ORSAY
n° d'ordre : 3383

UNIVERSITE DE PARIS-SUD

CENTRE D'ORSAY

THESE

présentée

Pour obtenir

Le TITRE de DOCTEUR 3ème cycle

PAR

Mademoiselle HUET Sylvie

SUJET : L'ANALYSE STATISTIQUE DES DOSAGES RADIOIMMUNOLOGIQUES

soutenue le 14 Février 1983 devant la Commission d'examen

MM. DACUNHA-CASTELLE Didier Président

JOLIVET Emmanuel

MARTINET Jack

TERQUI Michel

Ce document doit être remis à l'étudiant à l'issue de la soutenance.

ORSAY
n° d'ordre : 3383

UNIVERSITE DE PARIS-SUD

CENTRE D'ORSAY

THESE

présentée

Pour obtenir

Le TITRE de DOCTEUR 3ème cycle

PAR

Mademoiselle HUET Sylvie

SUJET : L'ANALYSE STATISTIQUE DES DOSAGES RADIOIMMUNOLOGIQUES

soutenue le 14 Février 1983 devant la Commission d'examen

MM. DACUNHA-CASTELLE Didier Président

JOLIVET Emmanuel

MARTINET Jack

TERQUI Michel

Ce document doit être remis à l'étudiant à l'issue de la soutenance.

Il me faut avant tout remercier toutes les personnes qui m'ont aidée à réaliser ce travail

L'équipe du Laboratoire de Biométrie du C.N.R.Z. à Jouy-en-Josas, et tout particulièrement : Monsieur JOLIVET, pour tout le temps qu'il a consacré à m'écouter et me conseiller ; Messieurs MESSEAN et VILA pour leurs conseils précieux concernant l'étude du modèle non-linéaire ; Mademoiselle PICART pour la frappe du texte.

Mesdames CAILLOL, SOLARI, VERMEIRE, PALY, Messieurs TERQUI et MARTINET, chercheurs ou techniciens à l'I.N.R.A., qui m'ont fourni les données traitées dans cette thèse, et qui m'ont permis grâce à leur patience et à leur gentillesse de comprendre dans le détail le principe d'un dosage radioimmunologique.

Monsieur DACUNHA-CASTELLE qui m'a aidée à mener à bien ce travail.

-§-§-§-§-

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41543524_0002

SOMMAIRE

-=-=-=-=-

INTRODUCTION

I- LES DOSAGES BIOLOGIQUES

I.1. But et implications statistiques

I.2. Notations

I.3. Le modèle

I.3.1. La relation dose-réponse
I.3.2. La condition de similitude

I.4. Les hypothèses statistiques

I.4.1. Le modèle
I.4.2. Conséquences de la condition de similitude pour deux modèles très simples

I.5. Transformation des données

I.5.1. Linéarisation
I.5.2. Hétérogénéité de la variance
I.5.3. Attention aux conséquences de la transformation

I.6. Estimation d'un intervalle de confiance. Théorème de Fieller

I.6.1. Dosage à courbes parallèles
I.6.2. Dosage à rapport de pentes

I.7. Plan d'expérience

I.7.1. Dosage à courbes parallèles
I.7.2. Dosage à rapport de pentes

ANNEXE : Le théorème de Fieller

II- LES DOSAGES RADIOIMMUNOLOGIQUES

II.1. Introduction

II.2. Notations et particularités

II.3. La courbe de réponse ou courbe standard

II.3.1. Les hyperboles

II.3.2. Les polynômes

II.3.3. Méthode tronquant les données

II.3.4. Le modèle logistique à deux paramètres

II.3.5. La méthode logit-log

II.3.6. Le modèle logistique à quatre paramètres

II.3.7. Les fonctions splines

II.4. Hétérogénéité de la variance - Pondération

II.5. Estimation de la courbe standard

II.5.1. Méthode logit-log

II.5.2. Modèle logistique à quatre paramètres

II.6. Estimation des doses inconnues et calcul d'un intervalle de confiance

II.6.1. Estimation

II.6.2. Calcul d'un intervalle de confiance à l'aide du théorème de Fieller

II.6.3. 2ème méthode, Méthode de calcul d'un intervalle de confiance proposée par Ross.

II.6.4. Comment se servir des dilutions ?

II.7. Plan d'expérience

II.7.1. Cas particulier : la régression linéaire

II.7.2. Le modèle logistique à quatre paramètres

II.8. Problème de dosages avec "extraction"

II.8.1. Exposé du problème

II.8.2. Conséquences sur la variabilité des réponses

II.8.3. Conséquences sur l'estimation des doses inconnues

ANNEXE A : Etude de la variabilité de la variance

ANNEXE B : Estimation des paramètres de la courbe standard et des préparations inconnues

ANNEXE C : Le programme DOSSTAT

ANNEXE D : Calcul d'une région de confiance pour une fonction des paramètres

CONCLUSION

BIBLIOGRAPHIE

-§-§-§-§-§-

R E S U M E

Les pratiques courantes pour effectuer l'analyse statistique des dosages radioimmunologiques reposent sur des hypothèses statistiques ignorées ou oubliées des utilisateurs, souvent fausses, rarement vérifiées. L'objet de ce travail est donc après avoir critiqué les méthodes usuelles, de proposer de nouvelles méthodes d'analyse. Souvent ces méthodes ont déjà été formulées par d'autres, mais parfois, il s'agit d'en développer de nouvelles qui permettent de mieux répondre aux questions posées par les biologistes.
Enfin, un programme conversationnel, d'utilisation très simple est proposé.

Mots-clefs : dosage biologique - dosage radioimmunologique - modèle linéaire - modèle non-linéaire.

INTRODUCTION

Le point de départ de cette recherche est un travail sur l'analyse statistique de dosages radioimmunologiques de la prolactine. Ce travail était demandé par Mrs KAAN et MARTINET, chercheurs au Laboratoire de Physiologie de la Lactation du C.N.R.Z. à Jouy-en-Josas. Deux questions étaient alors posées (HUET, S., 1978) :

- Quelles sont les limites d'utilisation de la courbe standard ?
- Peut-on diminuer la variabilité inter-dosages ?

Il est apparu très rapidement qu'il était difficile de donner une réponse à ces questions, sans auparavant remettre en cause les pratiques courantes des laboratoires pour effectuer l'analyse statistique des dosages radioimmunologiques. En effet, ces pratiques reposent sur des hypothèses statistiques ignorées ou oubliées par les utilisateurs, souvent fausses, rarement vérifiées.

Sur quelles hypothèses reposent un dosage ? Comment les vérifier ? Quelles en sont les conséquences sur les méthodes statistiques à utiliser ? En grande partie, les réponses à ces questions ne sont pas spécifiques aux dosages radioimmunologiques, mais concernent les dosages de façon plus générale. Aussi, un premier chapitre sur les dosages biologiques tente de faire le point sur ces questions, les dosages radioimmunologiques n'étant qu'un exemple de dosage.

Toutefois, le plan d'expérience d'un dosage radioimmunologique diffère de celui du dosage biologique et ceci nécessite donc une étude spécifique. C'est l'objet du chapitre II, qui commence par l'exposé et la critique des méthodes statistiques les plus couramment utilisées. Mais de la critique, doivent naître les propositions.

Dans un premier temps, il s'agit de reprendre des propositions déjà formulées :

- ajustement de la courbe standard à l'aide du modèle logistique à quatre paramètres
- calcul d'un intervalle de confiance pour les doses inconnues à l'aide du théorème de Fieller (FINNEY, 1977).
- estimation de la variabilité de la variance (RAAB, 1981).

Mais il s'agit aussi dans certains cas de proposer de nouvelles méthodes qui permettent de mieux répondre aux questions posées par les biologistes : l'application de la méthode de calcul d'un intervalle de confiance proposée

par ROSS (1978) aux doses inconnues d'un dosage radioimmunologique permet de détecter les limites d'utilisation de la courbe standard.

Pour de nombreux dosages hormonaux (les dosages des stéroïdes par exemple), il existe un problème important lié à la technique du dosage : il y a perte d'une certaine quantité d'hormone lors de son extraction de la solution à doser. Cette perte est estimée à l'aide d'un paramètre appelé récupération. Comment prendre en compte la variabilité de la récupération lors de l'estimation des doses inconnues ? A l'aide de calculs très simples, nous tentons d'apporter quelques réponses.

De même, comment choisir les doses de la courbe standard pour estimer au mieux les doses inconnues ? Dans le cas précis qui nous intéresse, cette étude n'avait encore jamais été faite.

Souvent, il existe un grand fossé entre la pratique courante (estimation de la courbe standard "à la main"), et les méthodes proposées. Ces méthodes font appel à la statistique du modèle non-linéaire, et donc pour leur application, il est nécessaire de disposer de logiciels plus ou moins sophistiqués. Aussi, était-il indispensable de proposer les outils informatiques qui devaient permettre l'utilisation de ces méthodes.

Le programme proposé répond à deux critères :

- il doit utiliser de bons algorithmes numériques fiables. De tels programmes existent déjà.
FINNEY (1977) pour l'estimation des doses inconnues et le calcul d'un intervalle de confiance à l'aide du théorème de Fieller ; RAAB et al. (1980) pour l'estimation de la variabilité de la variance ; de nombreux autres programmes implantables sur différents calculateurs existent (RODBARD et LEWALD (1970) par exemple).

- il doit regrouper ces différents calculs, permettre l'application de nouvelles méthodes (calcul d'un intervalle de confiance à l'aide de la seconde méthode du paragraphe II.6.3.), être très simple d'utilisation, offrir des résultats directement exploitables par l'utilisateur.

Face à toutes ces questions, l'objet de ce travail est de fournir aux biologistes, un outil devant leur permettre de mieux dominer l'analyse statistique des dosages radioimmunologiques.

LES DOSAGES BIOLOGIQUES

I.1. BUT ET IMPLICATIONS STATISTIQUES .-

Imaginons l'expérience suivante : deux préparations contenant des œstrogènes et ayant une action sur l'utérus sont injectées à des rats femelles. On mesure le poids de l'utérus des rates. Nous pouvons chercher à estimer les effets de ces préparations sur le poids de l'utérus. Nous faisons alors une analyse de la variance. Par contre, si nous utilisons les variations du poids de l'utérus pour estimer l'efficacité d'une des préparations par rapport à l'autre, nous faisons un dosage biologique. Le plus souvent, l'expérience implique un stimulus (une vitamine, une hormone, un fongicide) appliqué à un individu (ici le mot individu est pris au sens large : un animal, une parcelle de tissu d'un animal, une plante, etc...). La valeur du stimulus ou la dose, varie suivant un plan d'expérience défini par le chercheur. L'amplitude ou la fréquence de la réponse de l'individu au stimulus dépend de la dose. La relation entre la dose et la réponse permet de déduire l'efficacité d'une dose de la réponse qu'elle induit. En général, les dosages biologiques sont de type comparatif. Il s'agit alors d'estimer l'efficacité d'une préparation testée du stimulus par rapport à une préparation standard de ce même stimulus. Une préparation standard peut être une substance agréée internationalement ou bien, une substance parfaitement "connue" à l'intérieur d'un laboratoire.

On distingue généralement différents types de dosages :

- dosages directs : le but est de mesurer les doses des préparations standard et testées (appelées doses critiques) qui produisent une réponse spécifiée à l'avance par l'expérimentateur.

Par définition, l'efficacité de la préparation testée par rapport à la préparation standard est la quantité de préparation standard équivalente en effet à une unité de préparation testée. Le rapport de ces doses critiques estime l'efficacité. Pour pratiquer une technique directe, la réponse doit être administrée de telle sorte que la quantité exacte nécessaire pour produire la réponse puisse être notée. Les dosages directs ont un champ d'application

limité. Ils sont peu utilisés.

- dosages indirects : dans un dosage indirect, plusieurs doses sont administrées, chacune à plusieurs individus. Nous distinguons deux types de dosages indirects :

1. pour chaque dose administrée nous enregistrons si une réponse caractéristique (la mort par exemple) s'est produite ou non : c'est un dosage à réponse tout ou rien.

2. La réponse est la mesure d'une caractéristique de l'individu (poids, poids d'un organe particulier, temps de survie), c'est un dosage à réponse quantitative.

Le dosage quantitatif suppose l'existence d'une relation entre la dose et l'amplitude de la réponse. Cette relation étant connue pour les deux préparations, on estime alors les doses "équivalentes", ce qui permet d'estimer l'efficacité correspondante ainsi que sa précision.

Depuis longtemps, les chercheurs de discipline comme la pharmacologie, la toxicologie et d'autres, font des comparaisons numériques entre différentes substances ayant des activités biologiques similaires à l'aide de tests sur des tissus organiques, le plus souvent sans l'aide de statisticiens. Pourtant FINNEY (1977, 1978) donne d'excellentes raisons pour lesquelles les statistiques doivent servir l'étude des dosages biologiques :

- parce que les dosages dépendent du type et de l'amplitude des réactions biologiques qui varient d'un individu à l'autre, ils sont en eux-mêmes soumis à des variations aléatoires. Une étude statistique (en particulier l'étude des plans d'expérience) peut apporter une amélioration de la précision, en même temps qu'une économie des ressources biologiques.

- pour certains dosages les problèmes théoriques ou méthodologiques interviennent de façon cruciale (estimation linéaire ou non-linéaire, hétérogénéité de la variance). Il faut alors rechercher une solution exacte ou approchée.

I.2. NOTATIONS .-

Dans la mesure du possible, les lettres majuscules sont réservées aux variables aléatoires ; les lettres minuscules désignant toujours un nombre réel connu ou non. Ceci correspond aux notations habituellement utilisées dans les principaux ouvrages de statistiques élémentaires (LEFORT, 1967 ; DAGNELIE, 1969 ; SNEDECOR et COCHRAN, 1971).

Pour ne pas alourdir le texte, nous désignerons la préparation standard par la lettre S et la préparation à tester par la lettre T.

Définissons quelques valeurs seuil correspondant à des variables aléatoires de lois usuelles :

$\mathcal{T}(\alpha,\ell)$ est la valeur vérifiant :

$$\text{Proba}\ (X_\ell > \mathcal{T}(\alpha,\ell)) = \alpha/2$$

où X_ℓ est une variable aléatoire de loi de Student à ℓ degrés de liberté.

$\mathcal{F}(\alpha,\ell_1,\ell_2)$ est la valeur vérifiant

$$\text{Proba}\ (X_{\ell_1,\ell_2} > \mathcal{F}(\alpha,\ell_1,\ell_2)) = \alpha.$$

où X_{ℓ_1,ℓ_2} est une variable aléatoire de loi de Fisher à ℓ_1 et ℓ_2 degrés de liberté.

$\mathcal{E}(\alpha)$ est la valeur vérifiant

$$\text{Proba}\ (X > \mathcal{E}(\alpha)) = \alpha/2$$

où X est une variable aléatoire de loi Normale $\mathcal{E}(\alpha) = \mathcal{T}(\alpha,\infty)$.

I.3. LE MODELE .-

A partir de maintenant nous limitons notre étude aux dosages indirects à réponse quantitative, les dosages radioimmunologiques s'y rapportant.

I.3.1. La relation dose-réponse

Nous noterons z la dose d'un stimulus particulier reçue par un individu et U la réponse mesurée. L'espérance de la réponse d'un échantillon d'une population d'individus s'écrit $E(U/z) = \mu$.

L'amplitude de la réponse standard dépend de la dose au sens où la réponse moyenne pour une dose est déterminée par la dose. Ceci s'exprime à l'aide d'une équation de régression : $\mu = F(z)$

La fonction F doit être bijective.

Comment allons nous déduire l'efficacité de l'équation de régression ? Quelles hypothèses devrons-nous formuler ?

I.3.2. La condition de similitude

Pour une réponse moyenne μ, soit z^S et z^T les doses de S et T pour lesquelles la réponse est μ. L'efficacité de T relativement à S pour la réponse μ est z^S/z^T.

Supposons que nous soyons confrontés au problème suivant :

la préparation standard est constituée d'une substance chimique C qui produit la réponse que nous mesurons ; la préparation testée T est une dilution de S dans un certain rapport ρ : T est chimiquement identique à S. T est une simple dilution de C dans un diluant n'ayant aucun effet sur la réponse mesurée. Que vaut ρ ? ρ est exactement égal au rapport z^S/z^T qui est indépendant de μ.

Lors d'un dosage, la condition de similitude consiste à supposer que la réponse produite par T est exactement la même que si T était une simple dilution de S dans un rapport ρ. (Le terme dilution implique un rapport ρ inférieur à 1. Pour un rapport ρ supérieur à 1, il suffit d'employer le terme concentration). Si cette condition n'est pas vérifiée, l'efficacité de T relativement à S ne peut pas s'exprimer à l'aide d'une seule valeur ρ.

Aussi, nous allons immédiatement donner une écriture algébrique de la condition de similitude. Notons F^S et F^T les fonctions de régression pour les préparations standard et testée respectivement :

$$\mu^S = F^S(z)$$

$$\mu^T = F^T(z)$$

Doser T (la préparation testée) relativement à S (la préparation standard) suppose que :

1) il existe un nombre ρ, l'efficacité relative, tel que 1 unité (par unité de poids ou de volume) de T soit aussi efficace que ρ unités de S.

2) ce rapport ρ définit toutes les doses équivalentes (au sens où elles produisent le même effet de S et T).

Si ces deux conditions sont vérifiées, la réponse d'un individu à une dose z de T est exactement la même que si l'individu reçoit la dose ρz de S. Pour toutes les doses z on peut alors écrire :

$$F^T(z) = F^S(\rho z)$$

Si les mesures de la réponse pour S et T ne démentent pas l'équation ci-dessus, cela ne démontre pas que la même substance pour S et pour T produit à elle seule la réponse ; la condition est nécessaire et non suffisante.

De même, si les mesures pour S et T démentent l'équation ci-dessus, cela ne nous conduit pas forcément à remettre en cause la condition de similitude. Il se peut par exemple que les doses où la réponse est mesurée aient été mal choisies ou bien que les conditions de dosage aient été différentes pour les deux préparations, ou bien encore que les fonctions de réponse F^T et F^S aient été mal choisies. Seules, les connaissances biologiques ou chimiques de nous possédons sur la préparation à tester T nous permettent de faire ou non cette hypothèse. Si nous la savons fausse, les méthodes statistiques que nous exposons ci-après ne s'appliquent pas. CORNFIELD (1964) traite un exemple où la condition de similitude n'est pas vérifiée ; il étudie donc les variations de z^S/z^T en fonction de μ.

Dans tout l'exposé qui suit, nous nous en tiendrons au cas où la condition de similitude est vérifiée et nous allons nous intéresser à l'estimation de ρ.

I.4. LES HYPOTHESES STATISTIQUES .-

I.4.1. le modèle

Notre but est d'estimer l'efficacité et de connaître la précision de notre estimation. Pour obtenir ces informations sur ρ, nous allons devoir faire des hypothèses sur la loi de la variable aléatoire U et sur la forme de la fonction F. Nous les ferons à partir des observations et en tenant compte de la connaissance que nous avons du phénomène biologique.

La fonction de réponse F dépend de plusieurs paramètres $\theta_1, \theta_2, \ldots, \theta_p$.

F sera notée $F(z,\theta)$ où θ est le vecteur à p lignes

$$\begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_p \end{pmatrix}$$

Nous pouvons donc écrire $F^S(z) = F(z,\theta^S)$, $F^T(z) = F(z,\theta^T)$ comme une des conséquences de la condition de similitude. ρ dépendant de θ^S et θ^T, nous pouvons espérer déduire l'information que nous cherchons sur ρ de l'information que nous obtiendrons sur θ. Ainsi, nous nous ramenons au problème suivant : étant donné la relation $\mu = F(z,\theta)$, nous voulons obtenir des informations sur θ à l'aide de n couples (U_i,z_i), i variant de 1 à n, où U_i approche μ_i à une erreur E_i près : $U_i = \mu_i + E_i$ et $E(E_i/z_i) = 0$.

Le cas fondamental est celui où les diverses observations sont indépendantes, de précision constante (Var $E_i = \sigma^2$ pour tout i) et où F est une fonction linéaire en θ. Par exemple :

$$F(z,\theta) = \theta_1 + \theta_2 z$$

$$F(z,\theta) = \sum_{\ell=0}^{p-1} \theta_\ell \, z^\ell \qquad p \text{ étant connu}$$

$$F(z,\theta) = \sum_{\ell=0}^{p-1} \theta_\ell \, g_\ell(z) \text{ où p et les fonctions } g_\ell \text{ sont connus.}$$

De plus, si U est Normalement distribuée, son espérance et sa variance suffisent à décrire sa loi, donc la connaissance de θ et σ^2 nous permettent de décrire entièrement la loi de U.

Au stade où nous en sommes, une première question se pose : comment choisir la fonction de réponse ? Si nous connaissons parfaitement le phénomène biologique, nous pouvons écrire exactement la relation entre la dose et la réponse ; les paramètres θ_i à estimer ont alors un sens physique. Mais, l'estimation que nous ferons de θ sera d'autant moins précise que le nombre des paramètres θ_i sera grand et que la fonction de réponse sera une fonction compliquée de ces θ_i. Il se peut même que numériquement on ne soit pas en mesure

de déterminer les θ_i. Aussi, on préfère souvent chercher une fonction de réponse d'équation plus simple, réduisant le nombre de paramètres, même si ceux-ci n'ont plus aucun sens physique. Nous sommes alors confrontés à un problème de choix. Si notre connaissance du phénomène biologique n'est pas suffisante pour décrire la fonction de réponse, de fait le problème du choix de F nous est posé.

Les données observées sont-elles correctement ajustées ?

La procédure d'estimation des paramètres est-elle simple ?

Peut-elle conduire à de fausses solutions ? Sommes nous capables d'énoncer les propriétés des estimateurs des paramètres ? Sont-ils proches de la vraie valeur ? Les connait-on avec une bonne précision ? Nous ferons notre choix, en tantant de prendre en compte l'ensemble de ces critères.

I.4.2. Conséquences de la condition de similitude pour deux modèles très simples

Supposons que la relation entre la dose et la réponse soit linéaire : par exemple $F(z,\theta) = \theta_1 + \theta_2 z$. Pour toutes les doses z appartenant à l'intervalle d'étude, la condition de similitude entraîne : $F(z,\theta^T) = F(\rho z,\theta^S)$, ou encore $\theta_1^T + \theta_2^T z = \theta_1^S + \theta_2^S \rho z$.

Ainsi, les courbes de réponse pour S et T sont deux droites ayant même ordonnée à l'origine . Si β^S est la pente pour S et β^T la pente pour T, ρ vaut β^T/β^S. Un tel dosage est appelé dosage à rapport de pentes.

Supposons maintenant que la relation entre le logarithme de la dose et la réponse soit linéaire : $F(z,\theta) = \theta_1 + \theta_2 \ln z$. La condition de similitude entraîne alors : $\theta_1^T + \theta_2^T \ln z = \theta_1^S + \theta_2^S \ln\rho + \theta_2^S \ln z$

Les courbes de réponse pour S et T en fonction du logarithme de la dose sont deux droites parallèles. Si α^S et α^T sont les ordonnées à l'origine et β la pente, $\ln\rho$ vaut $(\alpha^T-\alpha^S)/\beta$. Nous parlerons alors de dosage à courbes parallèles.

Si nous supposons la réponse U Normalement distribuée, de variance constante σ^2, nous sommes alors ramenés au cas très simple de l'estimation des coefficients d'une droite par la méthode des moindres carrés. Nous en déduirons

une estimation de ρ par un simple calcul (TOMASSONE et al., 1982). De plus, le théorème de Fieller que nous verrons au paragraphe I.6. nous permet de calculer très facilement les limites de variation de ρ.

I.5. TRANSFORMATION DES DONNEES .-

Malheureusement les données que nous devons étudier ne se présentent pas toujours de façon si simple, et pour nous ramener à l'un des cas précédents nous cherchons à les transformer.

I.5.1. Linéarisation

Si la relation entre la dose et la réponse n'est pas linéaire, on peut chercher à transformer soit la dose, soit la réponse, soit les deux.

Notons x la dose transformée, nous restreindrons notre étude aux transformations suivantes :

$$x = z^{\lambda} \quad \lambda \neq 0$$

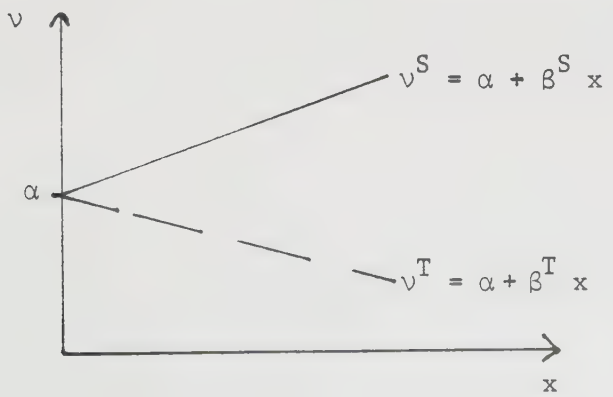
ou $$x = \ell n\, z \quad \text{si } \lambda = 0$$

Supposons que l'équation de régression $\mu = F(z,\theta)$ s'écrive $\mu = f(\theta_1 + \theta_2\, z)$ où f est une fonction connue strictement monotone. Nous supposons que λ a été choisi a priori ; ainsi seuls les paramètres θ_1 et θ_2 sont inconnus. Notons ν la réponse transformée définie par $\nu = f^{-1}(\mu)$.

Nous obtenons $\nu = \theta_1 + \theta_2\, z$

La condition de similitude impliquant que les mêmes transformations linéarisent à la fois les fonctions de régression de S et T, ρ se calculera de la même façon qu'au paragraphe I.4.2.

Si λ est non nul, x vaut z^{λ}, les courbes de la réponse transformée en fonction de x sont deux droites ayant même ordonnée à l'origine. Soit β^S et β^T leurs pentes, ρ vaut $(\beta^T/\beta^S)^{1/\lambda}$. Ce sont des dosages à rapport de pentes.



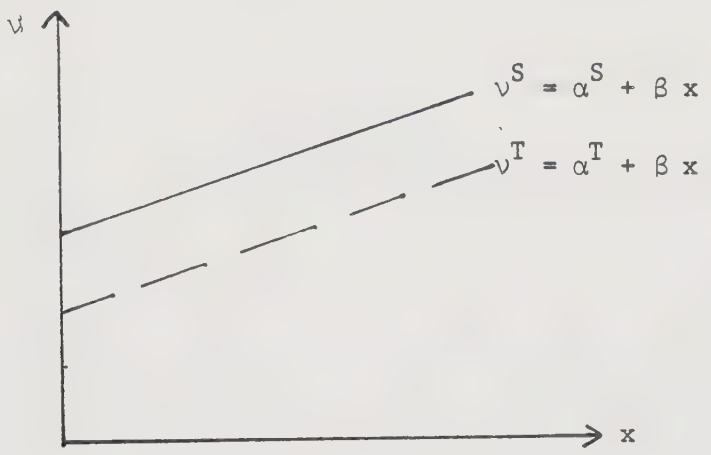
Dosage à rapport de pentes : courbes de réponse des préparations standard et testée après transformation de la dose ($x = z^{\lambda}$, $\lambda \neq 0$) et de la réponse.

—— : préparation standard.Droite d'équation $\nu = \alpha + \beta^S x$

--- : préparation testée. Droite d'équation $\nu = \alpha + \beta^T x$

D'après la condition de similitude, $\beta^T = \beta^S \rho^{\lambda}$

Si x vaut $\ln z$, les courbes de ν en fonction de x sont deux droites parallèles de pente β. Soit α^S et α^T les ordonnées à l'origine, $\ln\rho$ vaut $(\alpha^T - \alpha^S)/\beta$.



Dosage à courbes parallèles : courbes de réponse des préparations standard et testée après transformation logarithmique de la dose, et transformation de la réponse.

—— : préparation standard . Droite d'équation $\nu = \alpha^S + \beta x$

--- : préparation testée. Droite d'équation $\nu = \alpha^T + \beta x$

D'après la condition de similitude, $\alpha^T = \alpha^S + \beta \ln \rho$.

Les transformations les plus courantes de la réponse sont $\nu = \mu$; $\nu = \ln \mu$; $\nu = 1/\mu$; $\nu = \mu^{1/2}$.

Supposons que la transformation ait réellement permis de linéariser la courbe de réponse, si la variable aléatoire obtenue après transformation ($V = f^{-1}(U)$) est Normalement distribuée et de variance constante, comme au paragraphe I.4.2. nous sommes ramenés à l'estimation des coefficients d'une droite par la méthode des moindres carrés.

I.5.2. Hétérogénéité de la variance

Supposons que la variance de la variable aléatoire observée dépende de la dose ou de la réponse. Plus précisément soit $\sigma^2(U)$ la variance de U ($\sigma^2(U) = E(U - \mu)^2$), nous supposons que $\sigma^2(U)$ vaut $\phi(\mu)$ où ϕ est une fonction connue. Toujours dans le but de nous ramener au cas simple que nous savons traiter, nous allons chercher la transformation t telle que la variable aléatoire $U^* = t(U)$ soit de variance constante (TOMASSONE et al., 1982).

Exemples

- $\sigma^2(U) = \kappa\, \mu$ où κ est une constante

  $U^* = \kappa'\, U^{1/2}$ κ' étant une constante

- $\sigma^2(U) = \kappa\, \mu^2$

  $U^* = \kappa'\, \ell n\, U$

- supposons que la réponse soit comprise entre 0 et 1 et que la variance diminue au voisinage de ces bornes , $\sigma^2(U) = \kappa\, \mu(1-\mu)$. On obtient $U^* = \kappa\, \sin^{-1}\, (U^{1/2})$, la transformation angulaire bien connue appliquée aux pourcentages.

I.5.3. Attention aux conséquences de la transformation

Ainsi, nous pouvons être amenés à transformer les données, soit pour stabiliser la variance des erreurs, soit pour linéariser la relation entre la dose et la réponse. Nous espérons alors qu'après transformation ces deux buts soient satisfaits. C'est parfois le cas (quelques exemples sont exposés par KENDALL et STUART, 1962), mais aucune théorie ne nous permet de l'affirmer dans le cas général. En ce qui concerne les dosages biologiques, nous cherchons fréquemment à linéariser la relation entre la dose et la réponse ; les calculs de l'estimateur de l'efficacité et d'un intervalle de confiance deviennent alors très simples. Mais avant tout calcul, il nous faut vérifier que la linéarisation de la relation n'a pas introduit ou accentué l'hétérogénéité

de la variance.

De toutes façons, nous devons toujours formuler des hypothèses statistiques sur la variable que nous observons U, ou bien sur la variable obtenue après transformation V (comment varie l'espérance en fonction de la dose ?, Quelle est la loi de l'erreur ?).

Lorsque les observations que nous faisons le permettent, nous devons vérifier ces hypothèses. Nous possédons pour cela de nombreux outils appelés tests de validité.

. test de l'hypothèse de Normalité sur l'erreur des mesures de la réponse. Pour étudier les propriétés statistiques des estimateurs, nous devons faire des hypothèses sur la loi des erreurs des mesures de la réponse. Nous faisons très souvent l'hypothèse de Normalité sauf si on a de bonnes raisons pour supposer que la variable mesurée (ou transformée) suit une autre loi. PEARSON et HARTLEY (1970, page 67) exposent une méthode qui permet de tester l'écart des observations à l'hypothèse de Normalité. Malheureusement, le nombre de données récoltées lors d'un dosage biologique n'est en général pas suffisant pour permettre un tel test. Toutefois, nous sommes essentiellement intéressés par les lois des estimateurs et il n'est pas toujours nécessaire de faire des hypothèses très précises sur la loi des erreurs. Par exemple, supposons que pour i variant de 1 à n, la réponse U_i pour une dose z_i soit d'espérance $\theta_1 + \theta_2 z_i$

$$U_i = \theta_1 + \theta_2 z_i + E_i$$

Soit Θ_2 un estimateur de θ_2, Θ_2 est une fonction linéaire des réponses, plus précisément, Θ_2 est une fonction de la moyenne des réponses : $\sum_{i=1}^{n} U_i/n$. La loi de Θ_2 ne dépend donc que de la loi de $\sum_{i=1}^{n} U_i/n$ qui, sous des hypothèses très générales sur les U_i, peut être supposée Normale. (Voir par exemple, KENDALL et STUART, 1968, Tome 1, p. 193, l'énoncé du théorème central limite).

. test de l'hétérogénéité de la variance

1. Supposons que nous voulions comparer la variance des réponses pour les préparations standard et testée : pour S, la réponse U_i^S pour une dose z_i^S de

préparation standard est d'espérance μ_i^S et de variance σ^2 constante. Pour T, la réponse observée U_j^T pour une dose z_j^T de préparation testée est d'espérance μ_j^T et de variance τ^2 constante. Il s'agit, au vu des observations de U_i^S et U_i^T de tester l'hypothèse '$\sigma^2=\tau^2$' contre l'hypothèse '$\sigma^2\neq\tau^2$'. Pour effectuer un tel test, consulter par exemple DAGNELIE (1970, Vol. II, p. 50).

2. Supposons que nous voulions tester, l'homogénéité de la variance des réponses en fonction de la dose. U_{ij}, la jème répétition de la réponse en une dose z_i est supposée d'espérance μ_i et de variance σ_i^2. Il s'agit de tester l'hypothèse '$\sigma_i^2=\sigma^2$ pour tout i' contre l'hypothèse 'les σ_i^2 sont tous différents'. Le lecteur intéressé par un tel test pourra consulter DAGNELIE (1970, Vol. II, p. 53) ou bien PEARSON et HARTLEY (1970, Vol. I, p. 63). Si nous rejetons l'hypothèse d'égalité des variances, nous pouvons être conduits à estimer la variabilité de la variance (voir le paragraphe II.4.).

. <u>test d'écart au modèle ou test de l'hypothèse de linéarité</u>

Les observations présentent-elles un écart aux hypothèses faites sur la variation de l'espérance de la réponse en fonction de la dose ?

Restreignons d'abord notre étude au cas de la régression linéaire simple.

Supposons que la variable aléatoire U (il peut éventuellement s'agir de la variable transformée V) ait pour espérance une fonction linéaire de la dose ou d'une transformation de la dose. Nous observons n couples (U_{ij}, x_i) j variant de 1 à n_i, i variant de 1 à k et $n = \sum_{i=1}^{k} n_i$.

Le modèle est le suivant : $U_{ij} = \alpha + \beta x_i + E_{ij}$

où les erreurs E_{ij} sont supposées indépendantes, de lois gaussiennes, d'espérance nulle, de même variance σ^2 pour i variant de 1 à k. Notons μ_i l'espérance de U_{ij} : $\mu_i = E(U_{ij}/z_i)$. Il s'agit de tester l'hypothèse "$\mu_i = \alpha + \beta x_i$" contre l'hypothèse "$\mu_i \neq \alpha + \beta x_i$". Le principe du test repose sur la comparaison de la variabilité due au choix du modèle (c'est-à-dire à l'ajustement des données à une droite) à celle due à la variabilité des mesures (DAGNELIE, 1970, Vol. II, p. 276). En plus des hypothèses faites sur les erreurs ce test n'a de sens que si n>k>2, c'est-à-dire en présence de répétitions aux points x_i, sinon nous ne pourrions estimer la variabilité des mesures.

. comparaison de deux droites

Si à la fois pour la préparation standard et pour la préparation testée nous n'avons pas rejeté l'hypothèse de linéarité, suivant la transformation faite sur la dose, la condition de similitude impose aux deux droites soit d'être parallèle, soit d'avoir même ordonnée à l'origine. Si par exemple nous avons effectué une transformation logarithmique de la dose, le rejet de l'hypothèse de parallélisme des deux droites signifie que la condition de similitude n'est pas vérifiée (pour plus de rigueur, se reporter à la discussion du paragraphe I.3.2.). Il faudra alors modifier la technique du dosage, car dans ce cas, écrire que $\ell n\rho$ vaut $(\alpha_T-\alpha_S)/\beta$ n'a aucun sens. (pour effectuer ces tests voir TOMASSONE et al., 1982).

. analyse des résidus

Supposons que la réponse ait été observée une seule fois en chaque dose ce qui nous a empêché d'effectuer le test de linéarité, ou bien supposons qu'après avoir rejeté l'hypothèse de linéarité nous voulions en connaître la cause, l'étude des résidus peut nous aider à résoudre ces problèmes. (TOMASSONE et al. 1982).

Dans certains cas (ce sont les plus fréquents) nous ne pouvons pas nous ramener au modèle linéaire simple tel que nous venons de l'exposer. Par exemple, souvent la courbe de réponse n'est pas linéaire, et une transformation effectuée dans le but de linéariser la courbe de réponse introduit une forte hétérogénéité de la variance. Heureusement nous ne rejetons pas tous ces dosages. Des logiciels très performants et très facilement utilisables par de non informaticiens (par exemple MLP) permettent l'estimation de paramètres dans un modèle non linéaire et prennent en compte si cela est nécessaire l'hétérogénéité de la variance (VILA, 1976). Mais il n'est pas toujours indispensable de disposer d'un calculateur évolué et d'un logiciel "compliqué" pour faire une bonne analyse statistique. La méthode, dite itérative des moindres carrés pondérés, prend en compte la non-linéarité du modèle et nécessite peu de calcul (FINNEY, 1977).

I.6. ESTIMATION D'UN INTERVALLE DE CONFIANCE. THEOREME DE FIELLER .-

La réponse est la variable aléatoire U dont l'espérance pour une dose z est notée μ.

$$U = \mu + E$$

Nous observons n couples (U_i , z_i) : $U_i = \mu_i + E_i$. Les erreurs E_i sont supposées indépendantes, Normalement distribuées et de variance constante notée σ^2. Supposons que nous nous trouvons dans le cas simple où μ_i est une fonction linéaire de la dose ou d'une transformation de la dose.

I.6.1. Dosage à courbes parallèles

La relation entre le logarithme de la dose et la réponse est linéaire. Les deux relations dose-réponse s'écrivent alors :

$$\mu^S = \alpha^S + \beta\, x$$
$$\mu^T = \alpha^S + \beta \ln\rho + \beta\, x \quad \text{où } x = \ln z$$

équations de deux droites parallèles, séparées par la distance horizontale $\ln\rho$.

Soit $\bar{U}^S$ et $\bar{U}^T$, $\bar{x}^S$ et $\bar{x}^T$ les moyennes des réponses et des doses pour S et T ; soit B l'estimateur des moindres carrés de β. Si $\hat{U}^S$ et $\hat{U}^T$ sont les estimateurs des réponses en une dose z pour S et T, nous pouvons écrire :

$$\hat{U}^S = \bar{U}^S + B(x-\bar{x}^S)$$
$$\hat{U}^T = \bar{U}^T + B(x-\bar{x}^T) \qquad (*)$$

(*) *Supposons que nous ayons observé n^S couples (u_i^S, x_i^S) i variant de 1 à n^S et n^T couples (u_i^T, x_i^T) pour T, les deux droites étant parallèles, l'estimateur des moindres carrés de β est*

$$B = \frac{\sum_{i=1}^{n^S} (u_i^S-\bar{u}^S)(x_i^S-\bar{x}^S) + \sum_{i=1}^{n^T} (u_i^T-\bar{u}^T)(x_i^T-\bar{x}^T)}{\sum_{i=1}^{n^S} (x_i^S-\bar{x}^S)^2 + \sum_{i=1}^{n^T} (x_i^T-\bar{x}_i^T)^2}$$

où $\bar{u}^S$ et $\bar{u}^T$ sont les moyennes des réponses pour S et T.

La distance horizontale entre ces deux droites est

$$D = \bar{x}^S - \bar{x}^T + (\bar{U}^T - \bar{U}^S)/B$$

Un estimateur de ρ est R qui vérifie $D = \ell nR$. D est à une constante près, $\bar{x}^S - \bar{x}^T$ le rapport de deux fonctions linéaires des observations. Sous les hypothèses formulées ci-dessus sur U, $\bar{U}^T - \bar{U}^S$ et B sont Normalement distribuées, mais leur rapport n'est certainement pas Normal. Soit M le rapport $(\bar{U}^T - \bar{U}^S)/B$ et m le paramètre estimée par M, le théorème de Fieller (dont l'énoncé général et la démonstration sont donnés dans l'annexe) établit que les bornes de l'intervalle de confiance au niveau $1-\alpha$ de m, soit M_i et M_s sont

$$M_i \ , \ M_s = \left[M \pm \frac{\mathcal{C}\,\Sigma}{B} G^{1/2}\right] / (1-H)$$

où . Σ est l'estimateur de la variance résiduelle σ avec ℓ degrés de liberté (*)

. $\mathcal{C} = \mathcal{C}(\alpha, \ell)$

. $H = (\mathcal{C}^2\ \Sigma^2\ v_{22})/B^2$

. $G = v_{11} + M^2\ v_{22} - H\ v_{11}$

où $v_{11} = 1/n^S + 1/n^T$

$$v_{22} = 1/\left(\sum_{i=1}^{n^S} (x_i^S - \bar{x}^S)^2 + \sum_{i=1}^{n^T} (x_i^T - \bar{x}^T)^2\right)$$

Remarquons que si H est suffisamment petit devant 1 pour être négligé, la formule devient :

$$M_i \ , \ M_s = \left[M \pm \frac{\mathcal{C}\,\Sigma}{B} (v_{11} + M^2\ v_{22})^{1/2}\right]$$

L'approximation est bonne dès que H est inférieur à 0,05 ce qui exige que B soit neuf fois plus grand que son écart-type, si l'intervalle de confiance est calculé avec une probabilité de 95 %. Quand β n'est pas significativement différent de zéro, H devient plus grand que 1, et nous nous trouvons confrontés

---

(*) $\Sigma^2 = \left(\sum_{i=1}^{n^S} (U_i^S - \hat{U}_i^S)^2 + \sum_{i=1}^{n^T} (U_i^T - \hat{U}_i^T)^2\right) / (n^S + n^T - 3) \qquad \ell = n^S + n^T - 3$

à de nouvelles difficultés.

Nous ne donnerons pas de détails car en principe nous ne rencontrons que très rarement des dosages où H est plus grand que 1.

I.6.2. Dosage à rapport de pentes

Les deux relations dose-réponse s'écrivent :

$$\mu^S = \alpha + \beta^S x$$

$$\mu^T = \alpha + \beta^S \rho^\lambda x \qquad \text{où} \quad x=z^\lambda \ , \ \lambda \neq 0$$

équations de deux droites ayant même ordonnée à l'origine.

Soit $\beta^T = \beta^S \rho^\lambda$. Si B^S et B^T(*) sont des estimateurs de β^S et β^T, $R = (B^T/B^S)^{1/\lambda}$ estime l'efficacité relative ρ.

(*) *Estimer les coefficients des équations de régression indépendamment pour S et T ne serait pas correct. Nous devons tenir compte du fait que les deux droites se rejoignent à l'origine, comme nous avons tenu compte du parallélisme dans le paragraphe précédent. Nous nous contenterons ici de livrer les formules ; pour plus d'explications, consulter TOMASSONE et al. (1982).*

$$\text{Soit} \quad n = n^S + n^T \ , \quad SC^S_{xx} = \sum_{i=1}^{n^S} (x^S_i)^2 - \left(\frac{n^S \bar{x}^S}{n} \right)^2$$

$$SC^T_{xx} = \sum_{i=1}^{n^T} (x^T_i)^2 - \left(\frac{n^T \bar{x}^T}{n} \right)^2 \qquad SC^{ST}_{xx} = - \frac{n^S n^T}{n} \bar{x}^S \bar{x}^T$$

$$SC^S_{xU} = \sum_{i=1}^{n^S} x^S_i U^S_i - \frac{n^S \bar{x}^S}{n} (n^S \bar{U}^S + n^T \bar{U}^T)$$

$$SC^T_{xU} = \sum_{i=1}^{n^T} x^T_i U^T_i - \frac{n^T \bar{x}^T}{n} (n^S \bar{U}^S + n^T \bar{U}^T)$$

$$\Delta = SC^S_{xx} \, SC^T_{xx} - (SC^{ST}_{xx})^2$$

$$B^S = (SC^T_{xx} \, SC^S_{xU} - SC^{ST}_{xx} \, SC^T_{xU}) / \Delta$$

$$B^T = (-SC^{ST}_{xx} \, SC^S_{xU} + SC^S_{xx} \, SC^T_{xU}) / \Delta$$

Supposons $\lambda=1$; l'hypothèse n'est guère restrictive, si λ est différent de 1, nous estimons ρ^{λ} et calculons les limites de confiance pour ρ^{λ}, puis nous déduisons les résultats sur ρ en élevant à la puissance $1/\lambda$. Ainsi R vaut B^T/B^S. Le thoérème de Fieller établit que ρ appartient avec une probabilité $1-\alpha$ à l'intervalle dont les bornes R_i, R_s sont définies par

$$R_i, R_s = \left[R - H\frac{v_{12}}{v_{11}} \pm \frac{\mathcal{T}\Sigma}{B^S} G^{1/2}\right] / (1-H)$$

où

. $H = (\mathcal{T}^2 \Sigma^2 v_{11})/(B^S)^2$

. $G = v_{22} - 2R\, v_{12} + R^2 v_{11} - H(v_{22} - \frac{v_{12}^2}{v_{11}})$

. $\mathcal{T}$ et Σ ont même signification que dans le paragraphe I.5.1.

. $v_{11} = SC_{xx}^T / \Delta \quad v_{12} = SC_{xx}^{ST} / \Delta \quad v_{22} = SC_{xx}^S / \Delta$

où SC_{xx}^T , SC_{xx}^{ST} , SC_{xx}^S et Δ ont été définis précédemment.

Si H est suffisamment petit devant 1 pour être négligé, la formule devient

$$R_i, R_s = \left[R \pm \frac{\mathcal{T}\Sigma}{B^S} (v_{22} - 2R\, v_{12} + R^2 v_{11})^{1/2}\right]$$

Il n'est pas suffisant de donner une méthode de calcul d'un intervalle de confiance pour l'efficacité pour le seul cas très simple où la courbe de réponse est linéaire et la variance constante. Précisions toutefois que tout ce paragraphe I.6. peut être réécrit en remplaçant μ par ν et U par V après transformation des données telle que nous l'avons définie au paragraphe I.5.1. à condition que V soit supposée Normalement distribuée de variance constante σ^2. ELASHOFF (1981) montre à l'aide de simulations que dans le cas où la variance des réponses dépend de la dose ou bien dans le cas où les réponses pour deux doses successives sont corrélées, le théorème de Fieller conduit à sous-estimer la longueur de l'intervalle de confiance calculé pour l'efficacité. L'auteur propose alors d'appliquer la méthode du JACKKNIFE.

I.7. PLAN D'EXPERIENCE .-

Avant d'effectuer un dosage, interrogeons nous sur le choix du plan d'expérience. Combien de doses choisir pour S et pour T ? Combien de répétitions en chaque dose ? Quelle doit être l'étendue de l'intervalle de variation des doses pour S et T ? Sur quels critères allons-nous faire ce choix ? Bien sûr nous recherchons le "meilleur" plan d'expérience ! Mais "meilleur" ne veut rien dire si nous ne définissons pas les objectifs que nous voulons atteindre et la connaissance que nous avons déjà du phénomène biologique. Deux exemples très simples pour nous en convaincre.

Si nous cherchons à ajuster les réponses que nous avons observées en fonction de la dose à une droite, sans être certain a priori de la linéarité de la courbe de réponse sur l'intervalle des doses considéré, nous allons tester cette linéarité. Nous observons la réponse (ou la réponse transformée) V_{ij} en k doses x_i avec n_i répétitions en chacune de ces doses. Soit n le nombre total d'observations, $n = \sum_{i=1,k} n_i$, n étant fixé nous voulons déterminer k, x_1, n_1, x_2, n_2,..., x_k, n_k tel que le test de linéarité que nous allons effectuer soit le "meilleur" possible. Plus précisément si nous testons l'hypothèse $\mathcal{H}$: "la courbe est linéaire", contre l'alternative $\mathcal{A}$: "la courbe n'est pas linéaire", nous voulons que la probabilité de prendre la bonne décision, appelée puissance du test, soit la plus grande possible. Que nous dit le bon sens ? Donnons nous l'intervalle de variation des doses $[I_1, I_2]$. Nous choisirons évidemment $x_1 = I_1$, $x_k = I_k$. Si la courbe n'est pas une droite, nous devons être en mesure de mettre en évidence une courbure, ce que nous ferons d'autant mieux si nous choisissons un grand nombre de doses, bien réparties entre x_1 et x_k, soit k=n. Mais ce plan d'expérience est irréaliste car nous mesurons la réponse avec une certaine variabilité que nous devons estimer ce qui implique comme nous l'avons souligné au paragraphe I.5.3. que n>k>2. Il nous faut donc trouver un compromis entre un nombre important de doses pour mettre en évidence la courbure de la courbe de réponse et un nombre important de répétitions en chaque dose pour estimer au mieux la variabilité des réponses.

Nous énonçons ansi quelques principes concernant le plan d'expérience qui nous permettra de réaliser au mieux l'objectif que nous nous étions fixés : tester la linéarité de la courbe de réponse.

Changeons maintenant notre objectif : supposons que nous effectuons un dosage de routine et que nous savons que la courbe de réponse est linéaire. Notre nouvel objectif devient : estimer au mieux les paramètres de la régression linéaire. Quel est le plan d'expérience qui répondra à ce nouvel objectif ? Laissons l'intuition de côté et traitons cet exemple analytiquement.

Le modèle est le suivant :

$V_{ij} = \alpha + \beta x_i + E_{ij}$, les E_{ij} sont indépendants, Normalement distribués, d'espérance nulle, de variance σ^2.

Posons $\xi_i = x_i - x_.$, alors $V_{ij} = \alpha' + \beta \xi_i + E_{ij}$

Si A', B et Σ^2 sont les estimateurs de α' , β et σ^2,

$$A' = V_{..} \quad \text{où} \quad V_{..} = \sum_{i,j} V_{ij}/n$$

$$B = \sum_i n_i \xi_i (V_{i.} - V_{..}) \Big/ \sum_i n_i \xi_i^2 \qquad \text{où} \quad V_{i.} = \sum_j V_{ij} / n_i$$

$$\Sigma^2 = \mathcal{S}_H/(n-2) \quad \text{où} \quad \mathcal{S}_H = \sum_{i,j} (V_{ij} - A' - B \xi_i)^2$$

Notre but est d'estimer le mieux possible les paramètres α', β et σ^2. La somme des carrés $\mathcal{S}_H$ se décompose en une somme des carrés due à la variabilité des mesures et une somme des carrés due au choix du modèle, c'est-à-dire à l'ajustement des données à une droite. $\mathcal{S}_H = \sum_{i,j} (V_{ij} - V_{i.})^2 + \sum_i n_i (V_{i.} - A' - B \xi_i)^2$

Le choix du modèle étant supposé bon, la variance de $V_{i.}$ et $\sum_{i,j} (V_{ij} - V_{i.})^2$ sera d'autant plus faible que le nombre de répétitions en chaque i sera grand. Il faut deux points au moins pour estimer une droite, le meilleur plan d'expérience est donc celui où k=2, $n_1 = n/2$, $n_2 = n/2$. De plus, $\mathrm{Var}(A') = \sigma^2/n$, $\mathrm{Var}(B) = \sigma^2/(\xi_1^2 + \xi_2^2)$; nous diminuerons donc la variation

de B en augmentant $\xi_1^2 + \xi_2^2$ c'est-à-dire en choisissant l'intervalle des doses le plus étendu possible, $x_1 = I_1$, $x_2 = I_2$.

Ceux deux analyses très simples nous montrent que l'étude des plans d'expérience est un sujet très vaste. Plus précisément le choix du plan d'expérience dépend uniquement des objectifs que l'on se fixe. Revenons aux dosages biologiques. Quel est notre but ? Estimer l'efficacité et obtenir cette estimation avec le maximum de précision. Nous pouvons par exemple chercher les critères de minimisation de l'intervalle $[M_i, M_s]$ ou $[R_i, R_s]$, pour un nombre fixé d'individus $n=n^S+n^T$.

I.7.1. Dosage à courbes parallèles

Cherchons à minimiser le carré de la longueur de l'intervalle $m_s - m_i$, soit $\left(\frac{\Sigma \mathcal{T}}{B^S(1-H)}\right)^2 \{(1-H)\, v_{11} + (M - \bar{x}^S + \bar{x}^T)^2\, v_{22}\}$ (les notations ont été définies au paragraphe I.6.1.).

En regardant le sens de variation de la fonction définie par la formule précédente relativement à ses divers arguments, on constate que cette quantité est d'autant plus petite que :

. $\mathcal{T}$ est petit. La valeur de $\mathcal{T}$ est déterminée par le nombre de degrés de liberté de Σ^2. Il faut donc veiller à ce que le nombre d'individus soit maximum dans les limites techniques ou financières qui nous sont fixées.

. Σ est petit. La variance estimée par Σ^2 dépend de l'homogénéité des individus, d'un bon choix de la courbe de réponse et du soin avec lequel le dosage est fait.

. B est grand. B est l'estimateur de l'accroissement de la réponse transformée pour un accroissement d'une unité de la dose transformée. B estime la sensibilité des individus à un changement de dose.

. H est petit. Les différents points que nous venons de voir concourent à la minimisation de H. Nous retrouvons l'exemple que nous avons traité dans l'introduction du paragraphe I.6., mis à part le fait que B dépend des réponses des deux préparations S et T. Minimiser H revient à estimer β avec le maximum de précision. Nous ne pouvons apporter de réponse claire à cette question, cela dépend de la connaissance que nous avons a priori du phénomène biologique.

Malgré tout nous pouvons énoncer que nous devons rendre maximum $1/v_{22}$. Ceci sera d'autant mieux réalisé que l'intervalle de variation des doses sera le plus étendu possible (mais souvent cela entraîne une non-linéarité de la courbe de réponse).

. v_{11} est petit. $1/n^S+1/n^T$ est minimum pour $n^S=n^T$: les individus sont également répartis entre les deux préparations.

. $M-\bar{x}^S+\bar{x}^T$ est petit. Utiliser l'information a priori sur ρ, afin que les doses choisies soient telles que $\ell n\rho-\bar{x}^S+\bar{x}^T$ soit proche de zéro.

I.7.2. Dosage à rapport de pentes

Soit $I = (R_s-R_i)^2$ $\quad I = \left(\frac{\mathcal{C}\,\Sigma}{B^S(1-H)}\right)^2 \left[v_{22} - 2\,R\,v_{12} + R^2\,v_{11} - H\left(v_{22}-\frac{v_{12}^2}{v_{11}}\right)\right]$

Appelons x_M^S et x_M^T les doses maximum pour S et T, et divisons chaque dose de S et T par ces valeurs. Ainsi, pour les deux préparations la dose maximale vaut 1, ce qui élimine l'effet dû à l'unité des doses.

Alors, $\quad Var(B^S) = \sigma^2\, v_{11}/(x_M^S)^2$

$Var(B^T) = \sigma^2\, v_{22}/(x_M^T)^2$

$Cov(B^S,B^T) = \sigma^2\, v_{12}/x_M^S\, x_M^T$

où v_{11}, v_{12}, v_{22} sont définis au paragraphe I.5.2.

I devient $\quad I = \left(\frac{\mathcal{C}\,\Sigma}{B^S(1-H)}\right)^2 \left[\frac{v_{22}}{(x_M^T)^2} - \frac{2\,R\,v_{12}}{x_M^S\,x_M^T} + \frac{R^2\,v_{11}}{(x_M^S)^2} - \frac{H}{(x_M^T)^2}\,(v_{22}-\frac{v_{12}}{v_{11}})^2\right.$

où $\quad H = \left(\frac{\mathcal{C}\,\Sigma}{B^S\,x_M^S}\right)^2 v_{11}$

Nous pouvons encore écrire

$$I = \left(\frac{R\,\mathcal{C}\,\Sigma}{B^S\,x_M^S\,K(1-H)}\right)^2 \left[v_{22} - 2\,K\,v_{12} + K^2\,v_{11} - H\left(v_{22}-\frac{v_{12}^2}{v_{11}}\right)\right]$$

avec $K = R\,x_M^T/x_M^S$

Ainsi pour que I soit petit, il faut :

. $\mathcal{C}$ petit

. Σ petit

. $B^S x_M^S$ grand ; à la fois la pente de la courbe de la préparation standard, et la borne supérieure de l'intervalle de variation des doses pour S sont grands.

. H petit. Il reste donc à rendre v_{11} petit. Pour cela l'intervalle de variation des doses de S doit être large. On avait déjà x_M^S grand, il faut donc faire des mesures en des doses proches de zéro ou nulles.

. K grand. Nous avons déjà imposé à x_M^S d'être grand ; pour augmenter K, il nous reste à augmenter x_M^T, la valeur maximum des doses de T.

Bien sûr, augmenter x_M^S et x_M^T ne doit pas se faire au détriment de la linéarité de la courbe. En général, on choisit x_M^T proche de x_M^S/ρ (FINNEY, 1977) en utilisant l'information a priori sur ρ. Le cas idéal est celui où K est proche de 1, tout en lui étant légèrement inférieur.

. $[v_{22} - 2 K v_{12} + K^2 v_{11} - H(v_{22} - v_{12}^2/v_{11}]$ est petit.

Cette expression est minimisée en concentrant les individus aux doses extrêmes pour les deux préparations. Evidemment, la réalisation d'un dosage à l'aide d'un tel plan d'expérience implique que l'expérimentateur sait à l'avance que dans tout l'intervalle des doses, les courbes de réponse sont linéaires.

ANNEXE : LE THEOREME DE FIELLER

Enoncé

Soit A et B, deux fonctions linéaires d'une suite de variables aléatoires Normalement distribuées, de même variance σ^2.

Soit α, β, v_{11}, v_{22} et v_{12} définis de la façon suivante :

$E(A) = \alpha \quad E(B) = \beta \quad Var(A) = \sigma^2 v_{11} \quad Var(B) = \sigma^2 v_{22} \quad Cov(A,B) = \sigma^2 v_{12}$

Soit $\mu = \alpha/\beta$ et $M = A/B$. Soit Σ^2 un estimateur sans biais de σ^2 à ℓ degrés de liberté, indépendant de M. Soit M_i et M_s définis par

$$M_i, M_s = \left[M - H \frac{v_{12}}{v_{22}} \pm \frac{\tau \Sigma}{B} G^{1/2} \right] /(1-H)$$

où $$H = \frac{\tau^2 \Sigma^2 v_{22}}{B^2}$$

$$G = v_{11} - 2 M v_{12} + M^2 v_{22} - H(v_{11} - v_{12}^2/v_{22})$$

$$\tau = \tau(\alpha, \ell)$$

Alors, si H est inférieur à 1, la probabilité que μ soit compris entre M_i et M_s vaut $1-\alpha$. Si H est supérieur à 1, la probabilité que μ soit inférieur à M_i ou supérieur à M_s vaut $1-\alpha$.

Démonstration

Considérons la variable aléatoire $A-\mu B$. Son espérance est nulle, sa variance est égale à $\sigma^2(v_{11} - 2\mu v_{12} + \mu^2 v_{22})$. Nous pouvons affirmer que l'évènement $(A-\mu B)^2 \leq \tau^2 \sigma^2 (v_{11} - 2\mu v_{12} + \mu^2 v_{22})$ est réalisé avec une probabilité $1-\alpha$ si $\tau = \tau(\alpha, \ell)$.

Cette équation du second degré en μ admet deux racines M_i et M_s. Il est facile de vérifier que l'inégalité est vraie si μ est compris entre M_i et M_s quand H est inférieur à 1, ou bien si μ est inférieur à M_i ou supérieur à M_s quand H est supérieur à 1. Remarquons que supérieur à 1 est équivalent à b non significativement différent de zéro.

Application aux dosages biologiques

Supposons comme au paragraphe I.5.1. que l'équation de régression $\mu = F(z,\underset{\sim}{\theta})$ s'écrive $\mu = f(\theta_1+\theta_2\ x)$, f étant connue et strictement monotone.

1er cas

La réponse transformée $V = f^{-1}(U)$ est une variable aléatoire, d'espérance $\theta_1 + \theta_2\ x$ et de variance constante σ^2. Les observations de V sont indépendantes. Quelle que soit la transformation de la dose ($x = \ell nz$ ou $x = z^{\lambda}$), il suffit de consulter le paragraphe I.6. pour vérifier que le théorème de Fieller est applicable.

2ème cas

Les hypothèses statistiques sont faites sur la variable observée U. On suppose donc que U est une variable aléatoire d'espérance $f(\theta_1+\theta_2\ x)$ et de variance constante σ^2. Les observations de U sont indépendantes. Le modèle n'étant plus linéaire, les propriétés des estimateurs de $\underset{\sim}{\theta}$ (estimateurs sans biais, dont la variance ne dépend que de σ^2 et du plan d'expérience) ne sont plus vérifiées. Plus exactement on peut montrer qu'elles sont vraies pour de grandes valeurs de n. Dans la pratique, on fera toujours l'approximation : on supposera les hypothèses vérifiées et on appliquera le théorème de Fieller. Dans quelques cas particuliers nous serons amenés à discuter cette question plus à fond (voir le chapitre suivant sur les dosages radioimmunologiques).

LES DOSAGES RADIOIMMUNOLOGIQUES

II.1. INTRODUCTION .-

Les dosages radioimmunologiques sont très largement utilisés pour l'estimation de doses d'hormone. Ils permettent l'estimation de l'efficacité pour des quantités très faibles de préparations avec une précision réputée bonne. Toutefois, FINNEY (1977), RODBARD et FRAZIER (1975) critiquent les abus d'utilisation de la méthode. En effet, la précision semble si bonne, les courbes de réponse si régulières, qu'on est tenté d'ignorer les conditions de validité que nous avons exposées au chapitre I. Pour estimer la quantité d'hormone contenue dans une préparation inconnue, beaucoup utilisent une simple méthode graphique ; ou bien, pour estimer la courbe standard, ils effectuent une simple interpolation linéaire entre les points adjacents. Ces méthodes ne permettent pas d'estimer la précision autour des doses calculées ; de plus elles font appel à la subjectivité de l'expérimentateur, et entraînent une perte importante d'information concernant le système de dosage. Il faut raconnaître que la tentation d'utiliser une méthode graphique ou l'interpolation linéaire est forte. En effet, même si on connaît la fonction mathématique qui permettra d'ajuster les points de la courbe standard, les techniques d'estimation de ses paramètres peuvent rebuter ceux qui n'ont pas accès à des moyens de calcul suffisants ou ceux qui n'ont pas l'habitude de les utiliser. Cependant aujourd'hui cette difficulté doit être surmontée.

II.2. NOTATIONS ET PARTICULARITES .-

Le dosage radioimmunologique (ou RIA) d'une hormone repose sur la compétition entre une hormone radioactive et une hormone froide pour leur anticorps spécifique

$$H^* + H + A \rightleftharpoons AH^* + AH$$

où H^* représente l'hormone radioactive libre, H l'hormone froide libre, A l'anticorps libre, AH^* l'anticorps lié à l'hormone radioactive, AH l'anticorps lié à l'hormone froide.

La concentration d'anticorps étant volontairement limitée, et l'anticorps

se liant à la fois à l'hormone froide et à l'hormone marquée, l'addition d'hormone froide diminue la radioactivité du complexe "anticorps-hormone". Dans tous les tubes d'un dosage, on trouve la même concentration d'anticorps, la même concentration d'hormone radioactive, seules varient les quantités d'hormone froide. La réponse mesurée pour une dose z d'hormone est un comptage de radioactivité exprimé généralement en coups par minute ou cpm. On peut compter la radioactivité du complexe anticorps-hormone radioactive, notée B, ou F la radioactivité de l'hormone libre. Nous supposons dans la suite que B est mesuré.

SCHEMA D'UN DOSAGE RADIOIMMUNOLOGIQUE

1. Fixer les paramètres du dosage : $[H^*]$ et $[A]$ les concentrations d'hormone radioactive et d'anticorps qui seront ajoutées à chaque tube ; les conditions expérimentales.

2. La préparation standard

- Choisir le plan d'expérience

Choisir le nombre k de doses, choisir les doses et le nombre de répétitions de la mesure de la réponse en chaque dose, soit (z_1^S , n_1), $(z_2^S , n_2),\ldots, (z_k^S , n_k)$. Les z_i^S sont supposés parfaitement connus puisque ce sont différentes dilutions d'une hormone purifiée. Cette question sera partiellement étudiée au paragraphe II.7.

- mesurer les réponses

$(U_{1j}^S ; j=1,\ldots,n_1)$, $(U_{2j}^S , j=1,\ldots,n_2),\ldots, (U_{kj}^S , j=1,\ldots,n_k)$. Eventuellement, $z_1^S=0$ (respectivement $z_k^S = \infty$) ; les U_{1j}^S (respectivement les U_{kj}^S) sont alors n_1 (respectivement n_k) comptages pour une dose nulle (respectivement infinie) d'hormone froide du complexe anticorps-hormone radioactive, noté habituellement B_o (respectivement N).

3. Les préparations testées (appelées préparations inconnues)

Pour la préparation inconnue I, I variant de 1 à M,

- choisir le plan d'expérience

Choisir l, le nombre de dilutions de l'échantillon où la dose d'hormone est inconnue, les dilutions et le nombre de répétitions de la mesure de la réponse en chaque dilution, soit (d_1,r_1), $(d_2,r_2),\ldots,(d_\ell,r_\ell)$

- mesurer les réponses pour chacune de ces dilutions

$(U_{1j}^I , j=1,\ldots,r_1)$, $(U_{2j}^I , j=1,\ldots,r_2),\ldots, (U_{1j}^I , j=1,\ldots,r_1)$.

4. Effectuer l'analyse statistique

- choisir un modèle, c'est-à-dire : caractériser la courbe de réponse et étudier la variabilité des réponses. Ceci est l'objet des paragraphes II.3. et II.4.

- pour un modèle paramétrique particulier (le modèle logistique à 4 paramètres), une méthode d'estimation des paramètres est exposée dans le paragraphe II.5.

- En déduire l'estimation de l'efficacité de chacune des préparations inconnues relativement à la préparation standard, c'est-à-dire l'estimation de la dose d'hormone contenue dans chaque préparation inconnue (paragraphe II.6.).

1er exemple : Comptages observés pour un dosage radioimmunologique de la prolactin
Ces données proviennent du laboratoire de Physiologie de la Lactation du C.N.R.Z. à Jouy-en-Josas.

Préparation standard

Dose en ng/5 µℓ	réponses	
0	4394	4337
0,015625	4327	4205
0,03125	4219	4106
0,0625	3872	3900
0,125	3187	3386
0,25	2919	2826
0,3	2409	2588
0,4	2266	2204
0,5	2075	2092
0,75	1521	1558
1	1309	1330
1,25	1227	1176
2,5	659	689
5	293	349

Préparations inconnues

Volume de l'échantillon en µℓ	réponses	
5	2900	3100
5	950	1050

Tableau 1.a

Pour cet exemple, k=13, n_i=2 quel que soit i.
La réponse pour une dose infinie d'hormone froide n'a pas été mesurée, car on suppose que N est nul.
Pour chaque préparation inconnue ℓ=1. Le choix très fréquent d'un tel plan

d'expérience a plusieurs implications importantes. On ne peut pas estimer de courbe de réponse pour les préparations inconnues et on ne peut donc pas vérifier statistiquement la validité de la condition de similitude qui signifie que la réponse radioactive à un volume v de préparation inconnue est exactement la même que la réponse à un poids ρv d'hormone purifiée, ρ restant constant quel que soit la valeur de v ; ceci implique que l'anticorps devant se lier à l'hormone dosée est bien spécifique de cette hormone, c'est-à-dire qu'il n'existe pas dans la préparation inconnue une autre hormone susceptible de se lier à l'anticorps.

L'analyse statistique consiste donc en l'estimation d'une seule courbe de réponse à l'aide des seules données de la préparation standard. La courbe de réponse estimée, appelée courbe standard, permet l'estimation des doses inconnues. Si $\mu^S = F(z)$, si μ^I est l'espérance de la réponse pour un volume v de la préparation inconnue I, la concentration inconnue est alors égale à $v\ F^{-1}(\mu^I)$, F^{-1} désignant la fonction inverse de F.

2ème exemple

Comptages observés pour la courbe standard d'un dosage radioimmunologique de la testostérone. Ces données proviennent de la station de Physiologie Animale du C.N.R.Z. à Jouy-en-Josas.

Dose en pg/100 µℓ	réponses			
0	1312	1303	1265	1341
5	1204	1161		
10	979	968		
20	493	476		
50	192	188		
100	121	116		
200	71	71		
500	49	47		
1000	58	58		
∞	36	36	39	

Tableau 1.b

3ème exemple : Comptages observés pour la courbe stantard d'un dosage radioimmunologique du cortisol. Ces données proviennent du laboratoire de Physiologie de la Lactation du C.N.R.Z. à Jouy-en-Josas.

Dose en ng/.1 ml	réponses							
0	2868	2785	2849	2805	2779	2588	2701	275
0,02	2615	2651	2506	2498				
0,04	2474	2573	2378	2494				
0,06	2152	2307	2101	2216				
0,08	2114	2052	2016	2030				
0,1	1862	1935	1800	1871				
0,2	1364	1412	1377	1304				
0,4	910	919	855	875				
0,6	702	701	689	696				
0,8	586	596	561	562				
1	501	495	478	493				
1,5	392	358	399	394				
2	330	351	343	333				
4	250	261	244	242				
∞	131	135	134	133				

Tableau 1.c

II.3. LA COURBE DE REPONSE OU COURBE STANDARD .-

On peut chercher à prédire la forme de la courbe de réponse à partir de la loi d'action de masse. Considérons par exemple un cas très simple : le dosage de la prolactine.

Nous supposons que la fraction d'hormone radioactive liée à l'anticorps pour une dose infinie d'hormone froide est nulle (N=0). Nous allons chercher à exprimer le rapport B/Bo en fonction de z où, rappelons-le, B est la fraction d'hormone radioactive liée à l'anticorps en présence d'une quantité z d'hormone et Bo est la fraction d'hormone radioactive liée à l'anticorps en présence d'une quantité nulle d'hormone froide.

Précisons encore quelques notations :

On a $$H + A \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AH$$

si v_1 désigne la vitesse de formation de AH et v_2 la vitesse de la réaction inverse, k_1 et k_2 sont définies par :

$$v_1 = k_1 ([A] + [H])$$

$$v_2 = k_2 [AH]$$

où $[A]$ désigne la concentration d'anticorps libre, $[H]$ la concentration d'hormone libre, $[AH]$ la concentration du complexe anticorps lié à l'hormone.

On définit alors k la constante de Michaelis par $k = k_2/k_1$. Cherchons d'abord à exprimer Bo en fonction des constantes expérimentales. Nous mettons en présence de l'hormone radioactive et de l'anticorps sans ajouter d'hormone froide

$$H^* + A \rightleftharpoons AH^*$$

Nous faisons l'hypothèse que l'anticorps et l'hormone ont même concentration molaire (même nombre de molécules-gramme par litre), ce qui nous permet d'écrire la relation découlant de la loi d'action de masse du premier ordre, vraie une fois l'équilibre atteint :

$$[A] \times [H^*] = k [AH^*] = k\ Bo \qquad (1)$$

Si z^* et a sont respectivement les concentrations d'hormone marquée et d'anticorps, la relation (1) devient :

$$(a-Bo)\ (z^*-Bo) = k\ Bo$$

De ce polynôme du second degré en Bo on tire :

$$2\ Bo = z^* + a + k - ((a+z^*+k)^2 - 4\ a\ z^*))^{1/2}$$

On vérifie aisément que Bo est positif et inférieur à a et z^*.

Cherchons maintenant à exprimer B en fonction des constantes expérimentales et de z.

$$A + H + H^* \rightleftharpoons AH + AH^*$$

A l'équilibre, nous pouvons écrire :

$$\frac{[AH]}{[H]} = \frac{[A]}{k} = \frac{[AH^*]}{[H^*]} \quad \text{ou bien}$$

$$[AH] \times [H^*] = [H] \times [AH^*]$$

$$[A] \times [H^*] = k\,[AH^*] = k\ B$$

Nous en déduisons :

$$[AH]\ (z^*-B) = (z-[AH])\ B$$

$$(a-[AH]-B)\ (z^*-B) = k\ B$$

d'où $$[AH] = z\ B/z^*$$

et $$(a\ z^*-B(z+z^*))\ (z^*-B) = k\ B\ z^*$$

En résolvant cette équation du second degré en B, on obtient comme solution acceptable :

$$2B = \frac{z^*}{z+z^*}\ (z + z^* + a + k - \{(z + z^* + a + k)^2 - 4a\ (z + z^*)\}^{1/2})$$

De même que pour Bo, on vérifie que cette quantité est positive et inférieure à a et z^* ; quand z est nul elle vaut Bo, quand z tend vers l'infini, elle tend vers 0.

$$\text{Ainsi} - \frac{B}{Bo} = \frac{z^*}{z+z^*} \; \frac{z+z^*+a+k - \{(z+z^*+a+k)^2 - 4a(z+z^*)\}^{1/2}}{z^* + a+k - \{(z^* +a+k)^2 - 4a\, z^* \}^{1/2}}$$

Cette relation entre B/Bo et z semble bien compliquée et peu utilisable dans la pratique. Toutefois les quantités a et z^* ne sont pas choisies arbitrairement. Nous voulons que de fortes variations de z entraînent de fortes variations de B ; ceci sera réalisé si la quantité d'hormone radioactive z^* est faible devant les quantités d'hormone à doser. De plus la concentration optimale d'anticorps requise est telle que la concentration d'hormone marquée liée à l'anticorps en l'absence d'hormone froide (Bø) soit égale à la concentration d'hormone radioactive libre (la quantité d'anticorps libre restant à l'équilibre est alors négligeable). Ainsi on peut supposer que k est très petit devant a, z^* et z. La relation entre B/Bo et z se simplifie alors considérablement et devient

$$B/Bo = \frac{z^*}{z+z^*}$$

Le calcul que nous venons de faire pour la prolactine est en général beaucoup plus compliqué pour d'autres hormones et ne fournit pas des résultats aussi simples ; souvent le nombre de paramètres inconnus à estimer est très important. Aussi préfère-t-on traduire la forme de la courbe standard à l'aide d'un modèle mathématique.

Nous distinguerons sept méthodes, généralement utilisées par les biologistes, que nous classons de la façon suivante :

1) après transformation de la variable de réponse, la courbe de réponse est supposée hyperbolique ou logistique, et elle est décrite à l'aide de deux paramètres ; ou bien elle est supposée polynomiale, le degré du polynôme restant à déterminer.

2) On s'intéresse seulement à l'intervalle des doses où la courbe est linéaire.

3) On s'intéresse à la variable de réponse observée et non à une transformation de la variable. Deux méthodes sont alors envisagées :

. une méthode paramétrique. La courbe de réponse est supposée logistique et est décrite à l'aide de quatre paramètres.

. une méthode non paramétrique : l'ajustement de la courbe de réponse à l'aide de fonctions splines.

II.3.1. Les hyperboles

Appelons $\bar{c}$ et $\bar{d}$ les moyennes des réponses observées pour les doses infinie et nulle. Soit $U' = (U-\bar{c})/(\bar{d}-\bar{c})$. (Dans les paragraphes II.3. et II.4., nous intéressant seulement à la courbe standard, pour alléger le texte nous omettrons dans les notations l'indice S). z_i désignant les doses de la préparation standard, exceptées les doses nulle et infinie, le graphe des valeurs observées pour U'_{ij} en fonction de la dose z_i (figures 1.a., 2.a, 3.a) suggère que la courbe de U' en fonction de z soit une hyperbole(*) Ainsi la courbe de 1/U' en fonction de z est une droite.

En général, on n'observe cette linéarité que sur un intervalle réduit des doses, l'étendue de cette linéarité dépendant du dosage (figure 1.b., 2.b.,3.b.). Toutefois les données présentent souvent de sérieux écarts au modèle. Il est de plus impératif d'utiliser une méthode pondérée de régression linéaire car la transformation inverse amplifie l'erreur pour les fortes doses comme on le constate en examinant les figures 1.b.,2.b.,3.b. (Nous pouvons approcher la variance de 1/U' par $Var(U')/U'^4$, consulter par exemple DAGNELIE, 1970, p. 185).

(*) *Si F est la quantité d'hormone radioactive libre, T la quantité d'hormone radioactive totale, $\bar{t}$ la moyenne des réponses observées, les courbes des observations $(U-\bar{c})/(\bar{t}-\bar{c})$, $(\bar{U}-\bar{c})/F$ en fonction de la dose présentent également une forme hyperbolique.*

Figure 1 : Dosage de la prolactine

Soit $U'_{ij} = (U_{ij}-\bar{c})/(\bar{d}-\bar{c})$, $\bar{c}=0$, $\bar{d}=4365$

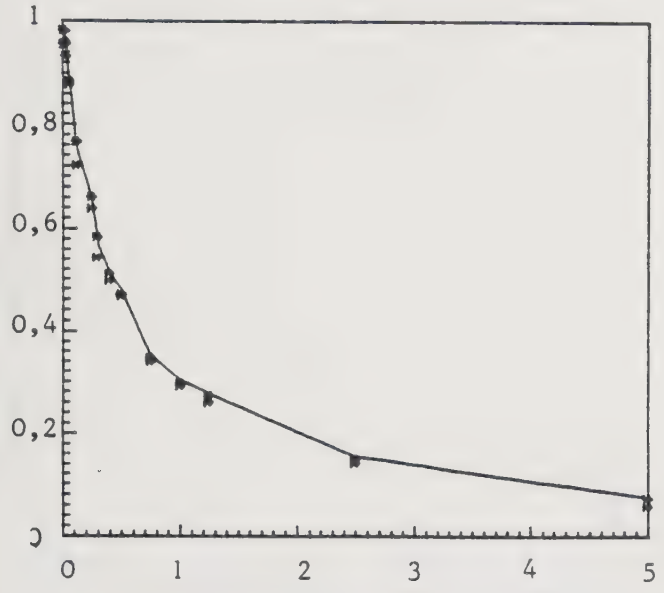


Figure 1.a.

* : valeur observée de U'_{ij} en fonction de z_i

— : segment joignant la moyenne des observations pour chaque i

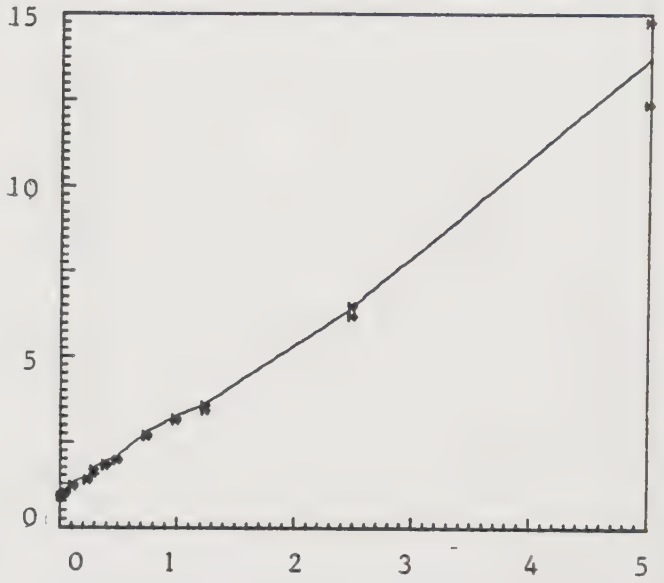


Figure 1.b.

* : valeur observée de $1/U'_{ij}$ en fonction de z_i

—— : segment joignant la moyenne de ces observations pour chaque i

Figure 2 : Dosage de la testostérone

$U'_{ij} = (U_{ij}-\bar{c})/(\bar{d}-\bar{c})$ $\bar{d}=1305$ $\bar{c}=37$

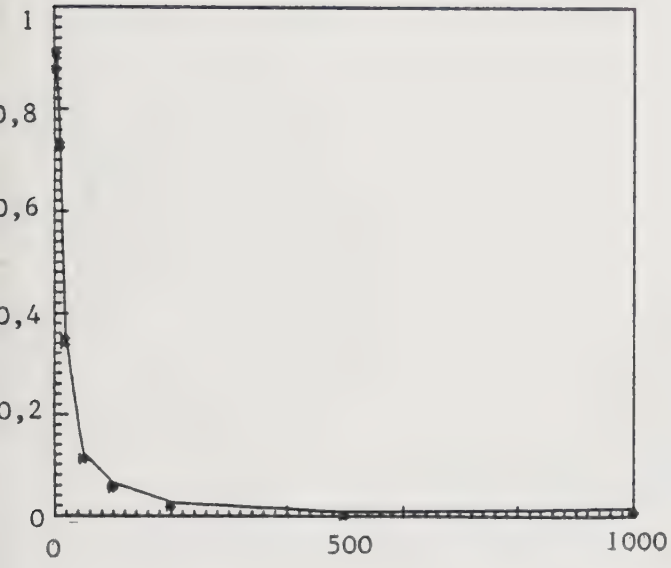


Figure 2.a.

* : valeur observée de U'_{ij} en fonction de z_i

—— : segment joignant la moyenne des observations

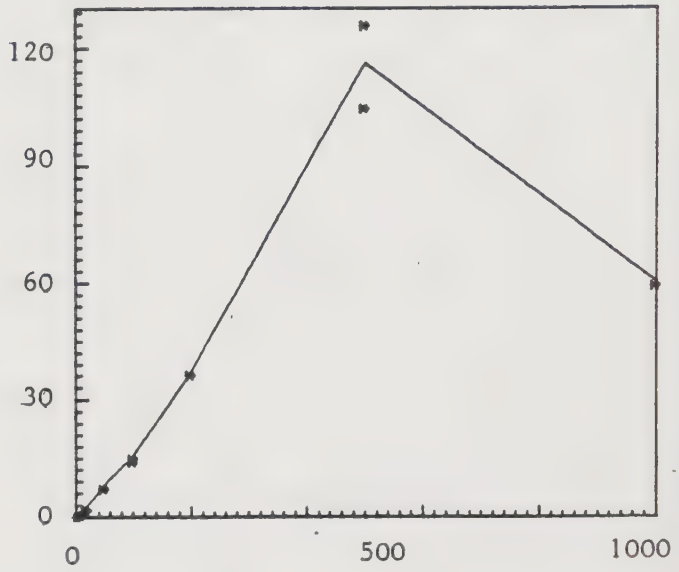


Figure 2.b.

* : valeur observée de $1/U'_{ij}$ en fonction de z_i

—— : segment joignant la moyenne de ces observations pour chaque i

Figure 3 : Dosage du cortisol

$$U'_{ij} = (U_{ij}-\bar{c})\,|\,(\bar{d}-\bar{c}) \qquad \bar{d}=2766 \qquad \bar{c}=133$$

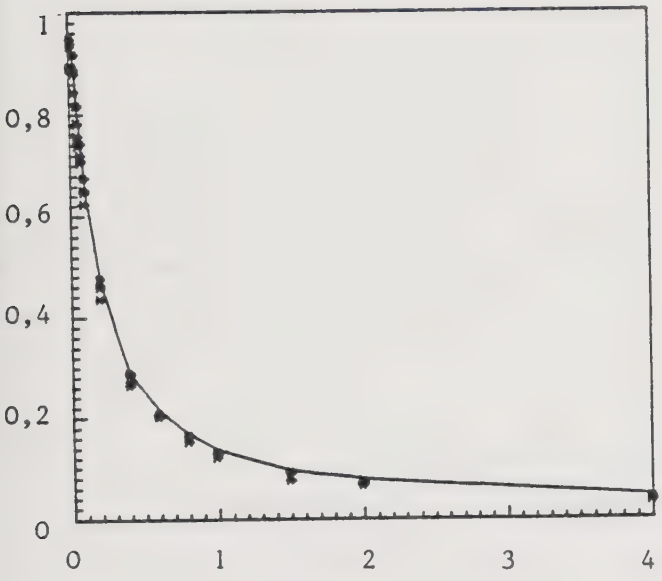


Figure 3.a.

* : valeur observée de U'_{ij} en fonction de z_i

— : segment joignant la moyenne des observations pour chaque i

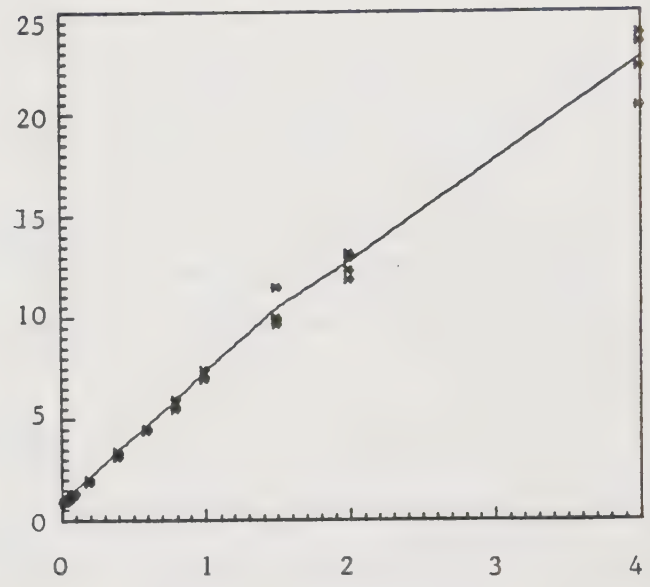


Figure 3.b.

* : valeur observée de $1\,|\,U'_{ij}$ en fonction de z_i

—— : segment joignant la moyenne de ces observations pour chaque i

II.3.2. Les polynômes

Certains ont cherché à ajuster aux réponses U'_{ij} un polynôme de degré p en x_i, où x_i est le logarithme de la dose. Si nous désignons par μ' l'espérance de U' la méthode consiste donc à estimer les coefficients $c_o, c_1, \ldots, c_p$ de la régression $\mu' = c_o + c_1 x + \ldots + c_p x^p$.

Bien que cette méthode soit générale elle présente plusieurs inconvénients et des études comparatives ont montré qu'elle n'était pas aussi performante que les méthodes logit-log ou logistique que nous verrons plus loin. Elle nécessite l'estimation de nombreux paramètres et l'inversion de la relation permettant l'estimation des doses inconnues peut s'avérer difficile à réaliser (dès que p est supérieur à 2, il est nécessaire d'utiliser une méthode numérique pour déterminer x en fonction de μ').

II.3.3. Méthode tronquant les données

Le graphe des valeurs observées U'_{ij} en fonction du logarithme de la dose, $x_i = \text{Log}(z_i)$, suggère que la courbe de U' en fonction de x soit approximativement linéaire sur une partie de l'intervalle de variation des doses (il suffit de ne pas considérer les doses extrêmes). Une telle méthode a été très largement utilisée par de nombreux chercheurs. Cependant, elle n'est pas toujours applicable (la figure 4 nous montre que pour le dosage de la testostérone, l'application de la méthode semble aventureuse, pour le dosage du cortisol, ce sont apparemment 2 droites qu'il faut ajuster et non une). De plus, cette méthode interdit l'estimation de doses en dehors de l'intervalle où la courbe est considérée linéaire.
A première vue, il semble raisonnable pour le dosage de la prolactine de limiter l'étude sur l'intervalle des doses [0,0625 ; 2,5].

Figure 4

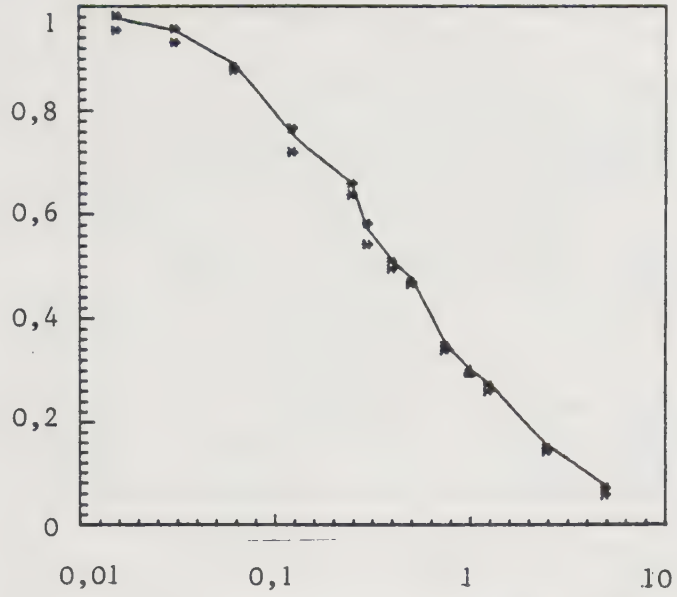


Figure 4.a. : Dosage de la prolactine

* : valeur observée de $U_{ij}^{\cdot}$, en fonction de x_i, $x_i=\text{Log}(z_i)$

——: segment joignant la moyenne des observations

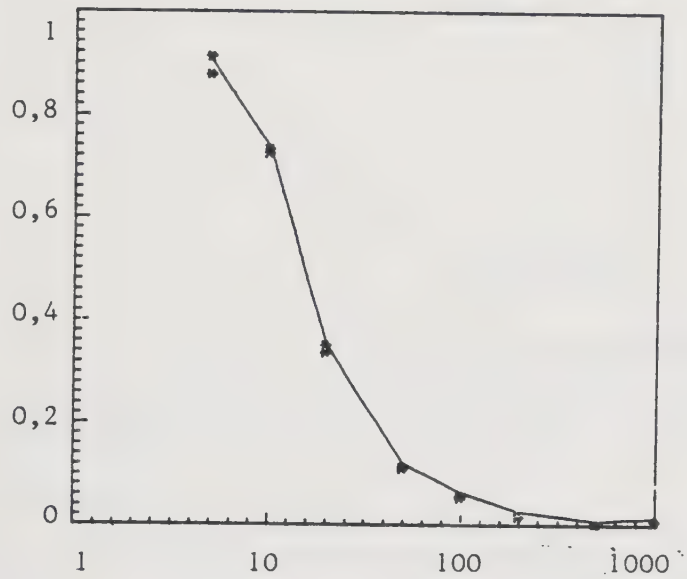


Figure 4.b. : Dosage de la testostérone

* : valeur observée de $U_{ij}^{\cdot}$, en fonction de x_i, $x_i=\text{Log}(z_i)$

——: segment joignant la moyenne des observations

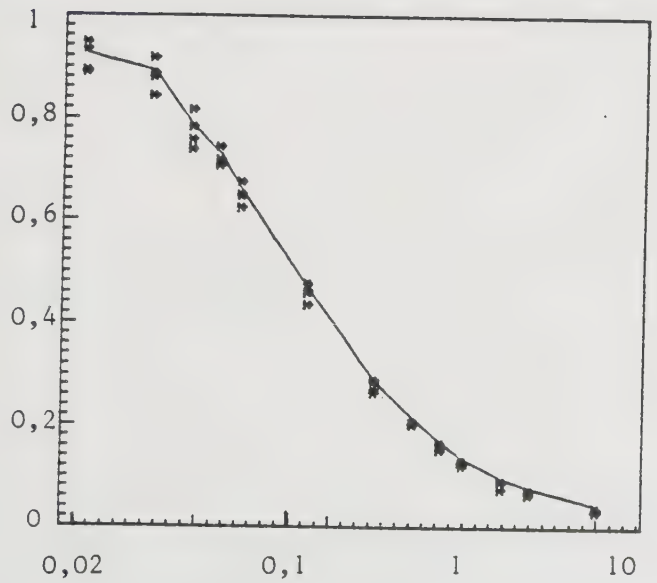


Figure 4.c. : Dosage du cortisol

* : valeur observée de $U_{ij}^{\cdot}$, en fonction de x_i, $x_i=\text{Log}(z_i)$

—— : segment joignant la moyenne des observations

II.3.4. Le modèle logistique à deux paramètres

Un examen de la figure 4 suggère que la courbe de U' en fonction du logarithme de la dose soit une sigmoïde.

Le modèle statistique est le suivant. $\bar{d}$ et $\bar{c}$ ayant été calculés à partir des observations de U aux doses nulle et infinie, nous supposons que la variable aléatoire observée est $U' = (U-\bar{c})/(\bar{d}-\bar{c})$. Si μ' est son espérance, $\mu' = f(\alpha+\beta x)$; nous allons chercher une fonction f qui traduise la forme sigmoïdale de la courbe de réponse. Parmi toutes les fonctions possibles, deux sont très largement utilisées :

- la fonction de répartition d'une gaussienne

$$f(\alpha+\beta x) = \int_{-\infty}^{\alpha+\beta x} \frac{1}{\sqrt{2\pi}} e^{-1/2t^2} dt$$

- la logistique

$$f(\alpha+\beta x) = \frac{1}{1+\exp-(\alpha+\beta x)}$$

Les deux fonctions ont un comportement similaire, (il peut y avoir une faible différence pour les valeurs de U' proches de 0 ou 1), nous restreignons donc dans la suite notre discussion à la logistique.

Remarques

1- Si la méthode hyperbolique s'applique avec succès, il en est de même pour la méthode du modèle logistique à 2 paramètres. β est alors égal à +1 (si x est le logarithme népérien de la dose) ou à +2,303 (si x est le logarithme en base 10 de la dose).

En effet, $1/(1+\exp-(\alpha+\beta x)) = 1/(1+e^{-\alpha} z^{-\beta})$ si $x = \ell nz$

ou $= e^{\alpha}/(e^{\alpha} + z^{-\beta})$

Il est bien clair que le modèle hyperbolique est un cas particulier du modèle logistique.

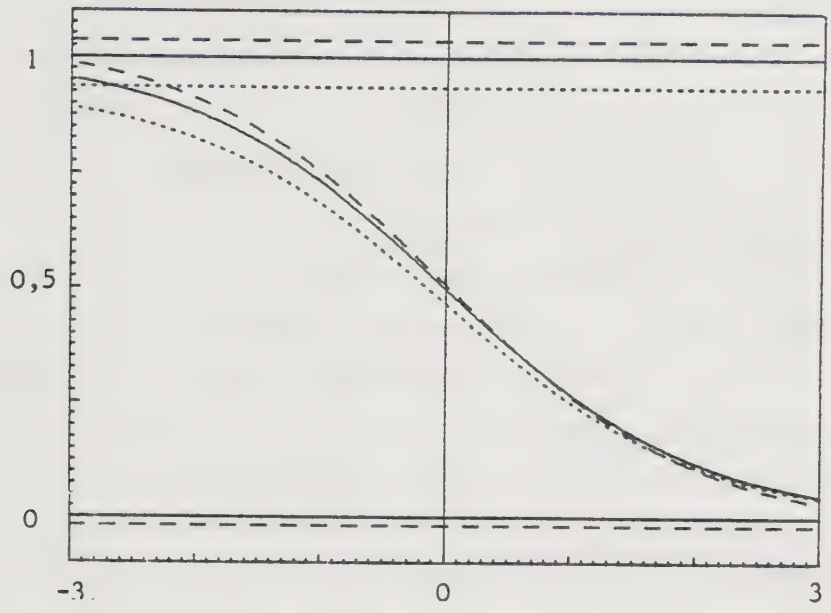
2- Les réponses pour les doses nulles et infinie ayant été estimées par $\bar{d}$ et $\bar{c}$, les valeurs ainsi obtenues sont considérées comme des constantes du modèle et la variabilité de ces estimations n'interviendra pas dans les calculs statistiques ultérieurs. D'autre part, toute erreur d'estimation importante sur ces paramètres entraîne un écart sérieux au modèle que nous allons tenter de décrire.

Notons V' la variable (U-c)/(d-c), d'espérance ν' où $\nu'=f(y)$ ($f(y)=1/(1+\exp -y)$). Soit $U' = \frac{d-c}{\bar{d}-\bar{c}} V' + \frac{c-\bar{c}}{\bar{d}-\bar{c}}$.

Si μ' est l'espérance de U', et si nous considérons que $\bar{c}$ et $\bar{d}$ sont des constantes $\mu' = \frac{d-c}{\bar{d}-\bar{c}} \nu' + \frac{c-\bar{c}}{\bar{d}-\bar{c}}$. Notons g(y) la fonction $\frac{d-c}{\bar{d}-\bar{c}} f(y) + \frac{c-\bar{c}}{\bar{d}-\bar{c}}$.

Simplement à l'aide d'un dessin, étudions les écarts entre f et g (figure 5).

Figure 5



—— Courbe de la fonction f(y) = 1/(1+exp-y) en fonction de y. f(y) est strictement comprise entre 0 et 1, elle décroît strictement entre les deux asymptotes f(y)=1, f(y)=0 ; elle admet un point d'inflexion en y=0, f(y)=1/2. Sur le dessin, les valeurs de y variant de -3 à 3, f(y) varie entre 0,05 et 0,95.

---- Courbe de la fonction $g(y) = \frac{d-c}{\bar{d}-\bar{c}} f(y) + \frac{c-\bar{c}}{\bar{d}-\bar{c}}$ en fonction de y. g(y) est strictement comprise entre $\frac{c-\bar{c}}{\bar{d}-\bar{c}}$ et $1 + \frac{d-\bar{d}}{\bar{d}-\bar{c}}$, elle décroît strictement entre les deux asymptotes $g(y) = 1 + \frac{d-\bar{d}}{\bar{d}-\bar{c}}$ et $g(y) = \frac{c-\bar{c}}{\bar{d}-\bar{c}}$, elle admet un point d'inflexion en y=0, $g(y) = \frac{d+c-2\bar{c}}{2(\bar{d}-\bar{c})}$, elle n'est pas symétrique.

Sur le dessin : d=3000, c=100, $\bar{d}$=2900 ($\bar{d}$ sous-estime d de 3,3 %), $\bar{c}$=150 ($\bar{c}$ surestime c de 5 %).

------ Courbe de g(y) en fonction de y pour c=$\bar{c}$=100, d=3000, $\bar{d}$=3200 ($\bar{d}$ surestime d de 6,7 %).

Il est donc indispensable d'estimer avec le maximum de précision Bo et N, par exemple en augmentant le nombre de répétitions des réponses aux doses zéro et infinie.

II.3.5. La méthode logit-log

Nous avons au paragraphe précédent déterminé une fonction f telle que $\mu' = f(\alpha+\beta x)$. Une fois les hypothèses statistiques faites sur la variable aléatoire U', les paramètres sont estimés à l'aide par exemple d'un logiciel d'estimation non-linéaire des paramètres (MLP).

La méthode logit-log consiste à estimer les paramètres α et β dans le modèle obtenu après transformation. C'est-à-dire, les hypothèses statistiques ne portent plus sur la variable U', mais sur la variable $V = \text{logit}(U')$ ou $V = \log \frac{U'}{1-U'}$ (figure 6). Ainsi, nous supposons V normalement distribuée, d'espérance $\nu = \alpha+\beta x$..Or les hypothèses que nous faisons sur V ne se déduisent aucunement de celles que nous avons faites précédemment sur U'. En particulier, étudions la variation de l'espérance de V en fonction de l'espérance de U'.

La comparaison de ν à logit (μ') nous conduit aux remarques suivantes :

1- ν vaut exactement logit (μ') pour μ' égal à 1/2 c'est-à-dire si l'espérance de la réponse U est exactement égal à (c+d)/2. (*).

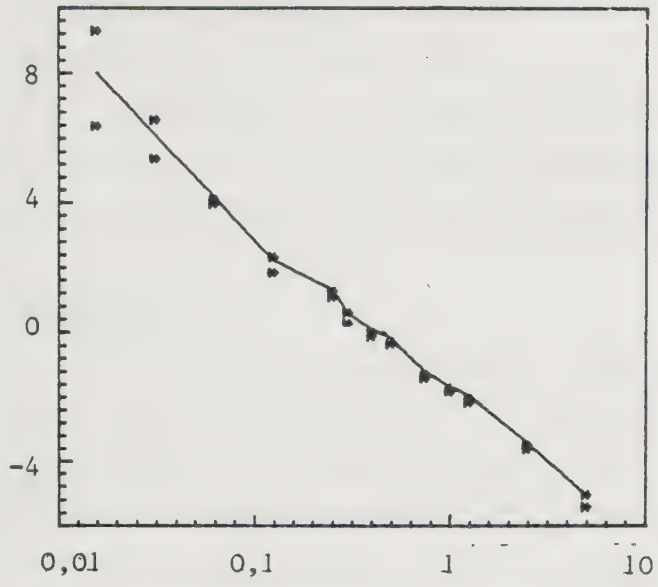
(*) *Notons $g(U')$, la fonction $g(U') = \text{logit}(U')$. A l'aide du développement en série de Taylor de la fonction g au voisinage de $U' = \mu'$, nous pouvons approcher $g(U')$ par* $g(\mu') + (U'-\mu')\, g'(\mu') + (U'-\mu')^2 \frac{g''(\mu')}{2}$.

Ainsi $\nu \approx g(\mu') + \frac{g''(\mu')}{2} E(U'-\mu')^2$

avec $g''(\mu') = -(1-2\mu')/(\mu'^2(1-\mu')^2)$.

$g''(\mu')$ *est nulle si* $\mu' = 1/2$.

Figure 6

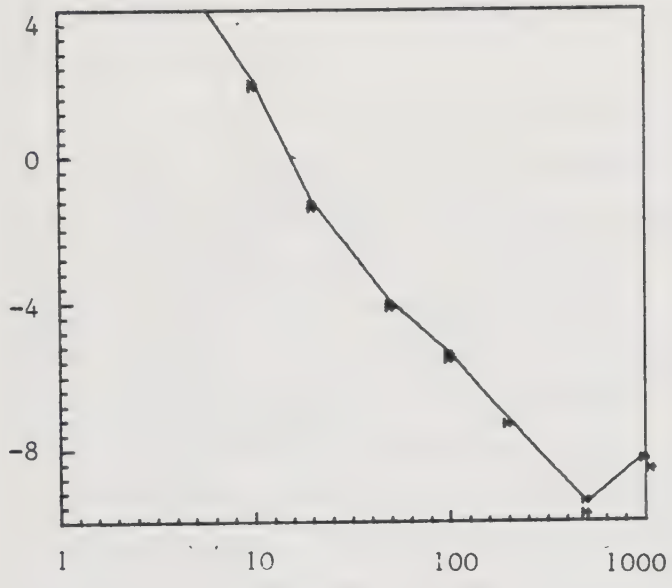


Soit V_{ij} = logit (U_{ij})

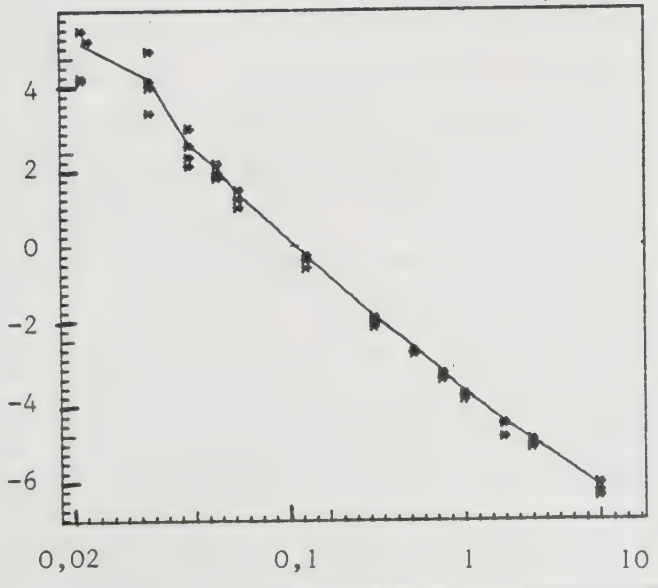
* : valeur observée de V_{ij} en fonction de x_i ; $x_i = \text{Log}(z_i)$

—— : segment joignant la moyenne des observations

6.a. : dosage de la prolactine



6.b. : dosage de la testostérone



6.c. : dosage du cortisol

2- Si μ' est différent de 1/2, ν s'éloigne d'autant plus de logit(μ') que μ' s'éloigne de 1/2. La qualité de l'approximation dépend de la variance de la réponse U'. (figure 7).

En pratique nous faisons toujours l'approximation et prenant en compte l'hétérogénéité de la variance induite par la transformation logit, une régression pondérée de V sur x fournit l'estimation des paramètres α et β.

Remarques

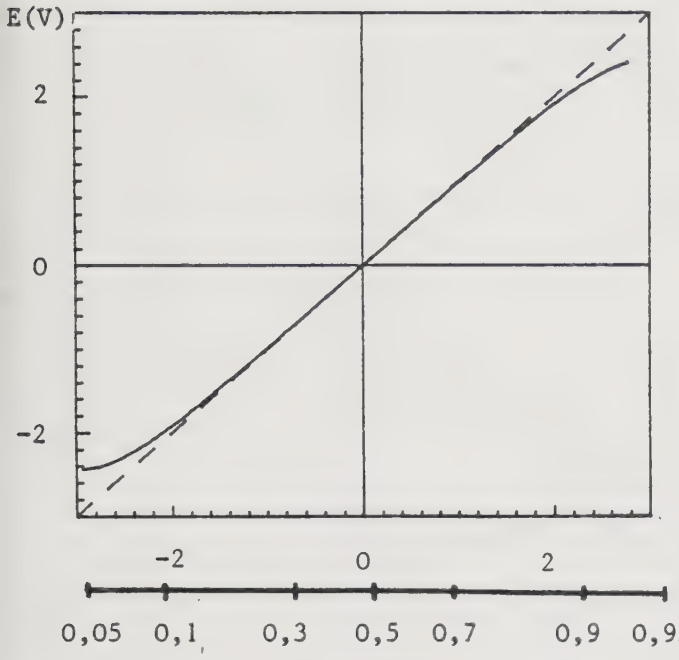
1. Toutes les remarques que nous avons faites au paragraphe II.3.4. pour la méthode logistique à 2 paramètres restent valables pour la méthode logit-log. En particulier pour la 2ème remarque concernant l'importance d'une bonne connaissance de $\bar{d}$ et $\bar{c}$ on se réfèrera à la figure 8.

2. Pour estimer α et β nous sommes amenés à éliminer les observations de U' supérieures à 1 ou négatives, faute de quoi nous ne pourrions calculer l'observation de V.

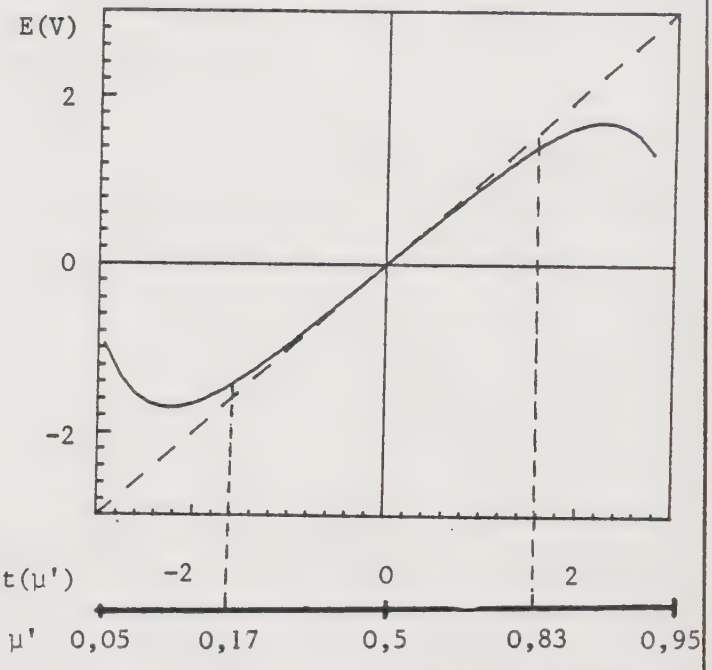
3. Pour estimer α et β certains emploient tout simplement une méthode graphique, d'autres utilisent la méthode des moindres carrés après pondération "à la main" des données transformées (élimination des données semblant aberrantes au vu du graphe des observations de V en fonction de x, limitation de l'analyse pour les doses dont la réponse observée de U' est comprise entre 0,2 et 0,8). Les estimations ainsi obtenues sont souvent proches de celles calculées après une régression pondérée, mais attention, d'une part ces méthodes d'estimation de α et β manquent d'objectivité, d'autre part, elles ne peuvent en aucun cas être appliquées pour un calcul ultérieur de la précision des estimateurs des doses inconnues. Toute analyse sans pondération alors que celle-ci est nécessaire conduit à une sous-estimation systématique de l'erreur sur une partie de la courbe et à une surestimation systématique sur l'autre. Si on prétend effectuer une analyse statistique sur laquelle on va s'appuyer pour confirmer ou non des hypothèses biologiques, il est impératif de chercher un modèle et d'énoncer des hypothèses statistiques qui s'adaptent aux observations si cela est possible, il est toujours déconseillé de manipuler les données, voire les éliminer pour arriver à ses fins : faire un calcul statistique.

Figure 7

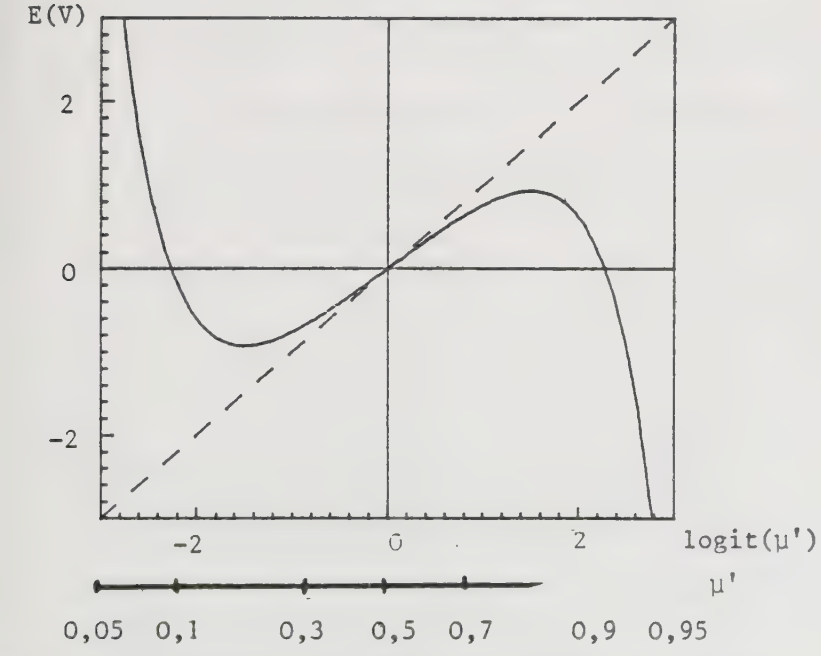
Soit U' une variable aléatoire d'espérance μ' et de variance σ^2. Soit V = logit(V'). Graphes pour différentes valeurs de σ^2 de E(V) en fonction de logit(μ') pour les valeurs de μ' comprises entre 5 % et 95 %, logit(μ') variant alors entre -3 et 3.



7.a. : σ^2=0,0025, correspond à une erreur de 10 % pour une réponse U' égale à 1/2. L'approximation de E(V) par logit(μ') est acceptable pour toutes les valeurs de U' considérées



7.b. : σ^2=0,01 correspond à une erreur de 20 % pour une réponse U' égale à 1/2. L'approximation de E(V) par logit(μ') est acceptable pour les valeurs de μ' approximativement comprises entre 17 % et 83 %.



7.c. : σ^2=0,04, l'approximation est rendue caduque.

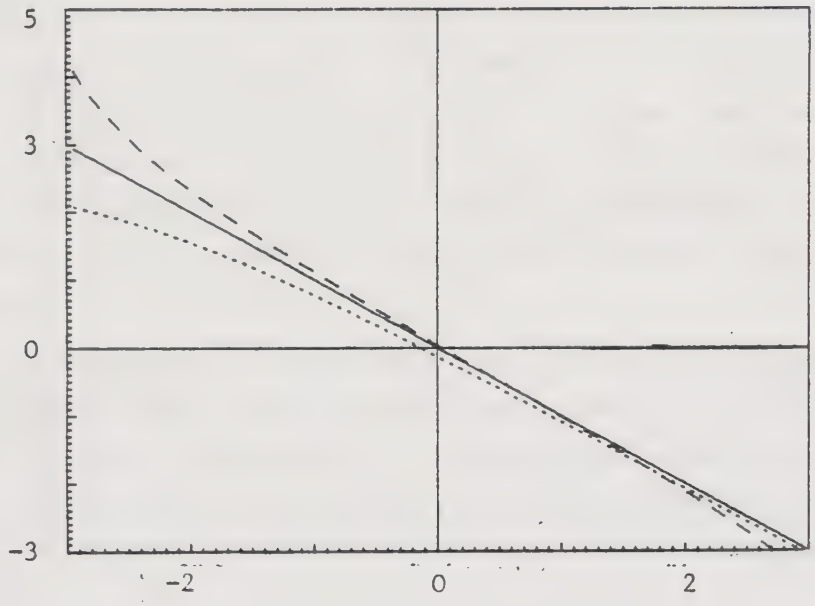
Figure 8

Soit $V' = (U-c)/(d-c)$ d'espérance $\nu' = f(\alpha+\beta x)$ $y = \text{logit}\ \nu' = \alpha + \beta x$

Soit $U' = (U-\bar{c})/(\bar{d}-\bar{c})$ $E(U') = \frac{d-c}{\bar{d}-\bar{c}}\ \nu' + \frac{c-\bar{c}}{\bar{d}-\bar{c}}$

Posons $\alpha=0$, $\beta=-1$ et faisons varier x entre -3 et 3.

Posons d=3000 , c=100



——— : courbe de y en fonction de x

- - - : courbe de logit(E(U')) en fonction de x pour $\bar{d}$=2900 , $\bar{c}$=150

----- : courbe de logit(E(U')) en fonction de x pour $\bar{d}$=3200 , $\bar{c}$=c

A l'évidence, ces deux dernières courbes présentent de sérieux écarts à la linéarité.

II.3.6. Le modèle logistique à 4 paramètres

Toutes les méthodes vues précédemment reposent sur des hypothèses statistiques faites sur les variables U' ou V obtenues après transformation de la variable observée U. Or, il est toujours préférable, lorsque cela est possible, d'effectuer une analyse statistique à partir des données observées et non pas à partir de données obtenues après transformation. En effet, nous formulons des hypothèses à l'aide de la connaissance que nous avons du phénomène observé. Bien sûr, dans ce paragraphe II.3., nous énumérons différents choix possibles sur la courbe de réponse de la préparation standard, et seules des considérations empiriques nous y ont amené. A première vue nous n'avons aucune raison de préférer un modèle à l'autre. Seules, des considérations d'ordre statistique nous permettrons de faire un choix et nous verrons par exemple dans le paragraphe II.4. que c'est la connaissance que nous avons de la variable U qui nous permet de faire des hypothèses sur la variabilité de sa variance.

Nous avons vu que la courbe de U' en fonction du logarithme de la dose est une sigmoïde, comprise entre les deux asymptotes 0 et 1. Ceci nous a permis d'introduire le modèle logistique à 2 paramètres. La courbe de U en fonction du logarithme de la dose est également une sigmoïde comprise entre les deux asymptotes c et d. D'où le modèle suivant appelé modèle logistique à 4 paramètres :

U, la réponse observée est une variable aléatoire d'espérance μ ; $E(U/z) = \mu$ et $\mu = c+(d-c)\ f(\alpha+\beta x)$ où $f(y) = 1/(1+\exp -2y)$.

Une procédure d'estimation itérative ou non linéaire permet l'estimation des paramètres c, d, α, β.

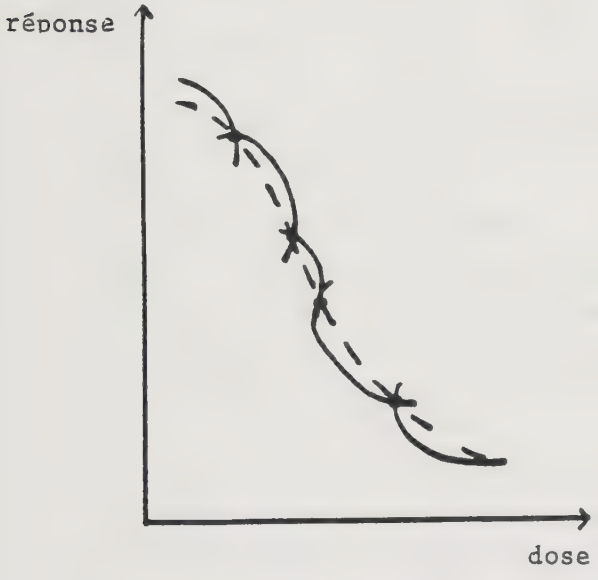
Remarques

1. Cette fois-ci, c et d sont des paramètres du modèle et il sera tenu compte de la variabilité de leurs estimateurs et de leur corrélation avec les autres estimateurs des paramètres du modèle α et β lors du calcul de la précision des doses inconnues estimées.

2. Précédemment, nous estimions d et c par $\bar{d}$ et $\bar{c}$ avec les seules réponses observées aux doses nulle et infinie. Avec cette méthode, toutes les réponses observées permettant l'estimation de la courbe standard contribuent à l'estimation de d et c. Nous pouvons ainsi espérer améliorer la précision des estimateurs de d et c.

II.3.7. Les fonctions splines

La méthode consiste à construire la courbe de réponse de la préparation standard, non pas à l'aide d'une seule fonction, mais à l'aide de plusieurs fonctions qui se joignent en des nœuds . Ces courbes sont des polynômes, de plus souvent de degré 3 (Pour plus de détails, voir par exemple WOLD, 1974).



plusieurs morceaux de courbe qui se joignent en des nœuds

Résultat final

Cette méthode est séduisante car il n'y a pratiquement aucune hypothèse à faire sur la forme de la courbe de réponse ; de plus, la fonction décrivant la courbe sur une partie de l'intervalle de variation des doses peut être complètement différente et donc indépendante de la fonction décrivant la courbe sur une autre partie de l'intervalle des doses. Remarquons également que la courbe de réponse est une fonction linéaire des paramètres à estimer ce qui peut considérablement simplifier les calculs que nous ferons ultérieurement.

Toutefois nous devons veiller à ce que la courbe de réponse reste strictement monotone, ce qui nous oblige déjà, en imposant des contraintes sur les paramètres à compliquer le modèle et à nous écarter des hypothèses du modèle linéaire. Ainsi on peut penser que la méthode des fonctions splines permet d'ajuster les données observées de toute courbe de réponse, y compris si les méthodes précédemment décrites échouent. Ceci est sûrement vrai, à condition de faire un choix convenable des nœuds . Certaines méthodes permettent donc l'estimation des coordonnées des nœuds en même temps que celle des coefficients des polynômes (JUPP, 1978). RAWLINS et YRJONEN (1978) ont écrit un programme qui permet l'ajustement de courbes de réponse des dosages radioimmunologiques à l'aide des fonctions splines.

II.4. HETEROGENEITE DE LA VARIANCE - PONDERATION .-

Quelle que soit la méthode utilisée pour l'estimation de la courbe standard, il nous faut avant tout calcul statistique poser des hypothèses sur la variabilité de la variable observée U.

La variance de la réponse U est-elle constante sur tout l'ensemble des doses étudié, $Var(U/z) = \sigma^2$. Ou bien, varie-t-elle en fonction de la dose ou de la réponse, $Var(U/z) = \sigma^2 \phi(\mu)$.

Comment ? Il est indispensable d'effectuer cette analyse pour les raisons que nous avons déjà évoquées lors de l'exposé de la méthode logit-log.

Nous pouvons nous poser la même question, non pas sur la variable observée U, mais sur les variables transformées U' ou V selon le modèle choisi. Toutefois, la méthode restant globalement identique et la variabilité de U' ou V se déduisant aisément de celle de U, nous exposons dans ce paragraphe l'analyse sur la variable observée U.

Si la question est ainsi posée, c'est que pour de nombreux dosages radioimmunologiques il a été constaté que la variance de U n'est pas homogène. En effet, U étant un comptage de radioactivité, on pourrait supposer que la loi de U en une dose z est une loi de Poisson de paramètre μ, et donc que sa variance est égale à μ. En fait, de nombreux facteurs d'erreurs intervenant lors du dosage, la variance de U n'est pas exactement égale à μ. L'hétérogénéité dépend de l'hormone, de la méthode de dosage, voire de l'expérimentateur. Aussi, nous estimerons l'hétérogénéité de la variance pour le dosage d'une hormone donnée dans un laboratoire donné, c'est-à-dire que dans ces conditions nous supposons qu'elle varie de la même façon d'un dosage à l'autre.

Revenons à la finalité de notre analyse. Nous cherchons un modèle permettant d'ajuster au mieux les données de la courbe standard, afin d'estimer le mieux possible les doses inconnues et les intervalles de confiance autour de ces doses. C'est l'estimateur de la dose inconnue et de sa précision qui nous interessent et nous sommes donc également amenés à formuler des hypothèses statistiques sur les réponses observées pour les doses inconnues. En particulier si la méthode de dosage est la même pour la préparation standard et les préparations inconnues, il est raisonnable de supposer a priori que la variabilité des réponses observées pour les préparations inconnues est la même que pour la préparation standard. Si les méthodes de dosages sont différentes

(par exemple; dosage direct pour la préparation standard, dosage avec extraction pour les préparations inconnues), il faudra s'assurer de cette similitude.

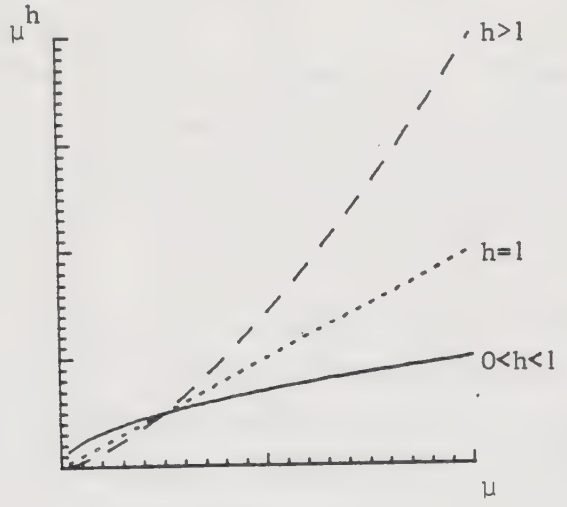
Pour décrire l'hétérogénéité de la variance des réponses d'un dosage radioimmunologique, deux modèles sont proposés :

$$\phi_h(\mu) = \sigma^2 \ \mu^h \qquad \text{(FINNEY, 1977)}$$

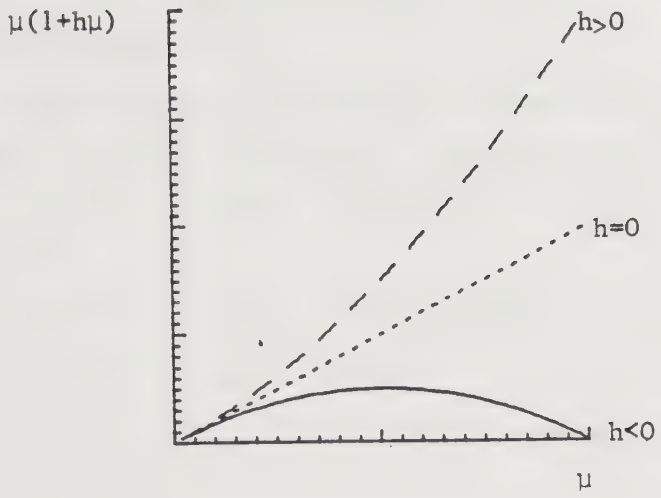
$$\phi_h(\mu) = \sigma^2 \ \mu \ (1+h\mu) \qquad \text{(RODBARD et al., 1976)}$$

Il s'agit après avoir choisi l'un des deux modèles d'estimer h.

Considérons le premier modèle proposé par FINNEY. Si h est positif la variance de la réponse croît avec la réponse, si h est nul, la variance est constante.



Considérons le second modèle proposé par RODBARD. h doit rester supérieur à $-1/\mu$ pour que $\mu(1+h\mu)$ reste positif. Si h est positif ou nul, la variance croît avec la réponse. Si h est négatif, la variance décroît pour les doses extrêmes.



h est appelé paramètre de second niveau par rapport aux paramètres décrivant la courbe standard, α, β, d et c. Ceci veut dire que l'on n'estimera pas simultanément les paramètres α, β, d, c et h de telle manière que les données observées soient au mieux ajustées, mais h étant connu a priori, on estimera α, β d et c. Nous chercherons donc h qui traduit au mieux l'hétérogénéité de la variance des réponses non pas d'un dosage donné, mais du dosage d'une hormone donnée, dans un laboratoire donné. Ceci pour plusieurs raisons :

1. Si nous considérons h comme un paramètre du modèle au même niveau que α, β, d, c, le modèle non linéaire ainsi défini devient complexe, et sa résolution numérique devient beaucoup plus compliquée, sans que cette complication nous apporte une meilleure compréhension du phénomène.

2. En général le plan d'expérience d'un R.I.A. est tel que seules deux réalisations de la réponse sont observées en chacune des doses. Aussi pour estimer h on a besoin d'un nombre important de données ; ceci est réalisé en mettant en commun les observations obtenues lors de différents dosages.

3. S'il est important de prendre en compte lors de l'analyse statistique du dosage la forme de la variabilité des réponses, il n'est pas nécessaire de connaître la valeur de h avec une grande précision. En effet, une légère modification de la valeur de h n'apporte pas de modification sur les estimations des paramètres de la courbe standard, de la dose inconnue et de sa précision.

Cette affirmation ne repose que sur de simples constatations (FINNEY et PHILLIPS, 1977) et il serait intéressant d'essayer de comprendre pourquoi il existe une telle robustesse.

Cherchons maintenant à estimer le paramètre de second niveau h.

Supposons que pour une dose z_i (connue ou non) nous notions U_{ij} la jème répétition de la réponse, j variant de 1 à k_i (en général $k_i=2$), i variant de 1 à k.

Supposons que $E(U_{ij})=\mu_i$ (nous ne faisons aucune hypothèse sur la relation entre μ_i et z_i) et $Var(U_{ij}) = \sigma^2 \phi_h(\mu_i)$

où $$\phi_h(\mu_i) = \mu_i^h \qquad (1)$$

ou $$\phi_h(\mu_i) = \mu_i(1+h\mu_i) \qquad (2)$$

Soit $U_{i.} = \frac{1}{k_i} \sum_j U_{ij}$ $\qquad S_i^2 = \sum_j (U_{ij}-U_{i.})^2$

RODBARD et al. (1976) ont proposé d'estimer h en régressant $\log(S_i^2/k_i-1)$ sur $\log U_{i.}$ pour le modèle (1), ou S_i^2/k_i-1 sur $U_{i.}$ et $U_{i.}^2$ pour le modèle (2). RAAB (1981) montre que cette méthode est peu efficace, c'est-à-dire que la variance de l'estimateur de h est grande. FINNEY et PHILLIPS (1977) ont estimé par la méthode du maximum de vraisemblance les k+2 paramètres $\mu_1, \mu_2, \ldots \mu_k$, h et σ^2. Or cette méthode n'est pas satisfaisante, car plus le nombre d'observations augmente plus le nombre de paramètres augmente. Les propriétés asymptotiques du maximum de vraisemblance ne sont alors plus vérifiées. RAAB (1981) présente une méthode dérivée du maximum de vraisemblance pour l'estimation de h, ainsi qu'un test permettant de tester l'adéquation au modèle. Un exposé succinct du principe de la méthode est fait en annexe de ce chapitre.

Remarquons que pour le premier modèle, la valeur de h ne dépend pas de l'intervalle de variation de la réponse. Plus précisément, si V=k U, h sera le même pour U et pour V pour le premier modèle, différent pour le second. Cette propriété peut nous conduire à choisir le premier modèle si par exemple

nous désirons comparer les valeurs de h pour des dosages faits dans différents laboratoires ou à l'aide de différentes méthodes.

1er exemple : Dosage de la prolactine

Nous disposons des données de 5 courbes standards de 5 dosages de la prolactine effectués dans les mêmes conditions au même laboratoire. Nous avons ainsi 92 réponses moyennes observées, avec 2 répétitions pour chacune ($k=92$, $k_i=2$, $i=1,\ldots,K$).

Un test d'homogénéité des variances non significatif nous permet de supposer que pour ce dosage la variance des réponses est constante. La figure 9 semble confirmer ce résultat.

Figure 9

Etude de la variabilité de la variance pour 5 courbes standards de 5 dosages de la prolactine. On observe U_{ij} pour j variant de 1 à 2, i variant de 1 à 92. Soit $S_i^2 = \sum_{j=1}^{2} (U_{ij}-U_{i.})^2$ et $U_{i.} = \sum_{j=1}^{2} U_{ij}/2$.

La figure 9-a représente la série des points observés pour S_i^2 en fonction des observations de $U_{i.}$.

Ces points ne semblent pas présenter une structure particulière.

Si nous supposons que $Var(U_{ij}) = \sigma^2 \mu_i^h$, alors $\ell n\ Var(U_{ij}) = \ell n\ \sigma^2 + h\ \ell n\ \mu_i$.

La figure 9-b représente la série des points observés pour $\ell n\ (S_i^2)$ en fonction des observations de $\ell n(U_{i.})$. Là encore le graphe ne met pas en évidence une telle relation entre les 2 variables.

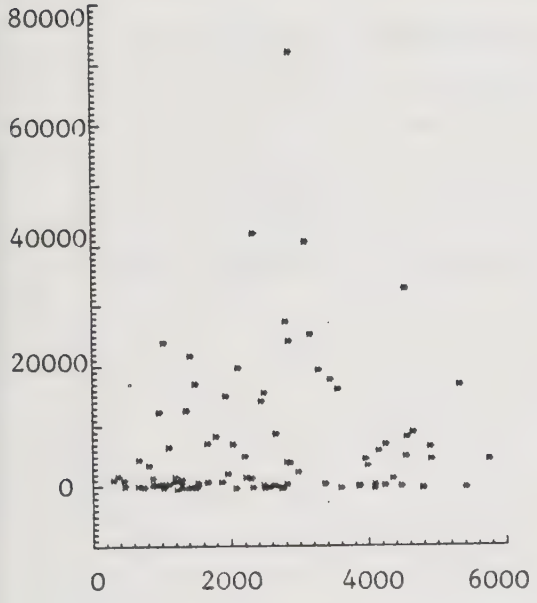


Figure 9.a.

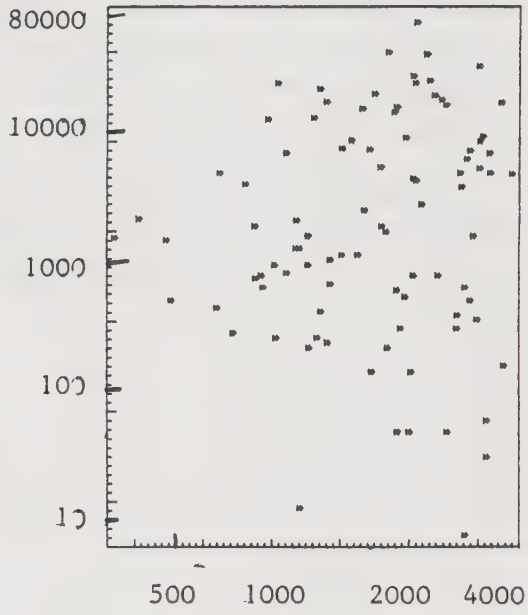


Figure 9.b.

2ème exemple : Dosage de la testostérone

Nous disposons de 10 courbes standards de 10 dosages de la testostérone effectués dans les mêmes conditions au même laboratoire. L'estimation des paramètres s'est faite avec 93 réponses moyennes observées, avec en général 2 répétitions en chaque point (parfois 3 ou 4). Sous l'hypothèse $\phi_h(\mu) = \mu^h$, h a été estimée par 1.26 avec un écart-type de 0,11. Le résultat du test d'adéquation au modèle nous permet de dire que cet ajustement convient. Sous l'hypothèse $\phi_h(\mu) = \mu(1+h\mu)$, h a été estimé par $9,4\ 10^{-5}$ avec un écart-type de $3\ 10^{-4}$. Le test d'adéquation au modèle nous mène à la même conclusion que précédemment. Mais dans le second cas, l'écart-type estimé autour de h étant très grand, nous préférons garder pour les calculs ultérieurs le premier modèle. Voir figure 10.

3ème exemple : Dosage du cortisol

Nous avons ici considéré toutes les réponses d'un dosage du cortisol, réponses pour la courbe standard, réponses pour les échantillons inconnus. L'estimation des paramètres s'est faite avec 130 réponses moyennes observées, avec 2 répétitions en chaque point pour les échantillons inconnus, 4 ou 8 pour les échantillons de la courbe standard. Sous l'hypothèse $\phi_h(\mu) = \mu^h$, h a été estimé par 1,7 avec un écart-type de 0,16. Sous l'hypothèse $\phi_h(\mu) = \mu(1+h\mu)$ h a été estimé par 0,0022 avec un écart-type de $55\ 10^{-5}$. Les tests d'adéquation nous permettent d'accepter les 2 modèles. (Voir figure 11).

Figure 10

Etude de la variabilité de la variance pour 10 courbes standards de 10 dosages de la prolactine. On observe U_{ij} pour j variant de 1 à k_i , i de 1 à 93 ; k_i vaut 2,3 ou 4.

Figure 10.a.: * série de points observés pour S_i^2 en fonction des observations de U_i.

—— courbe ajustée $\phi_h(\mu) = \mu^h$

--- courbe ajustée $\phi_h(\mu) = \mu(1+h\mu)$

Figure 10.b.: * série des points observés pour $\ell n(S_i^2)$ en fonction des observations de $\ell n\ U_i$.

—— courbe ajustée $\phi_h(\mu) = \mu^h$. Ce graphe met clairement en évidence la dépendance entre $\ell n(S_i^2)$ et $\ell n(U_i)$.

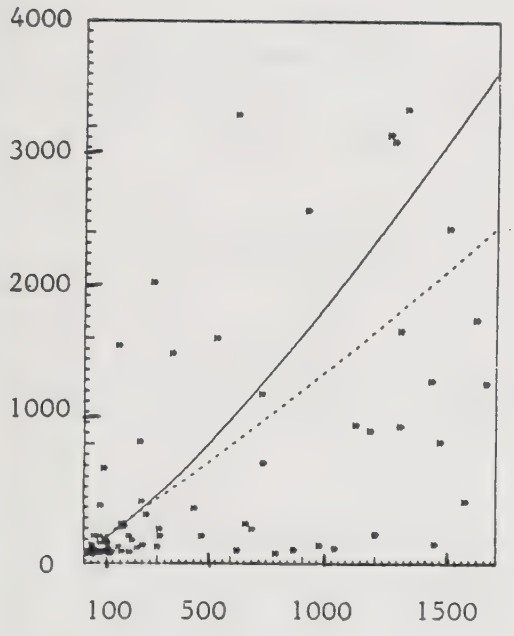


Figure 10.a.

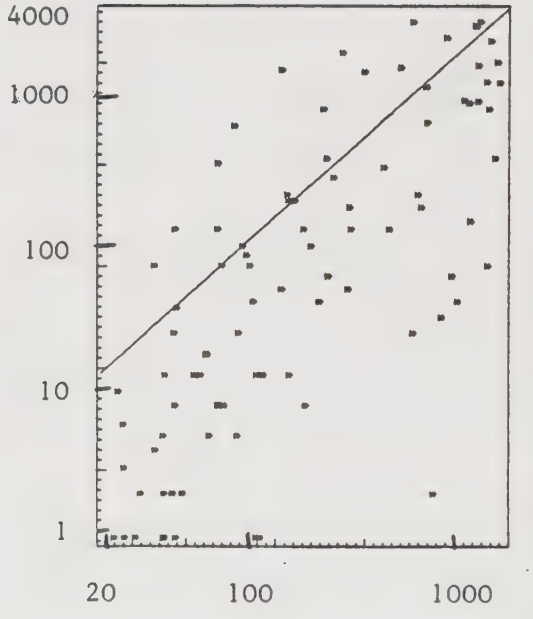


Figure 10.b.

Figure 11

Etude de la variabilité de la variance pour un dosage du cortisol. On observe U_{ij} pour j variant de 1 à k_i, i de 1 à 130.

k_i vaut 2,4 ou 8. Figures 11.a et 11.b. : mêmes légendes que figure 10.

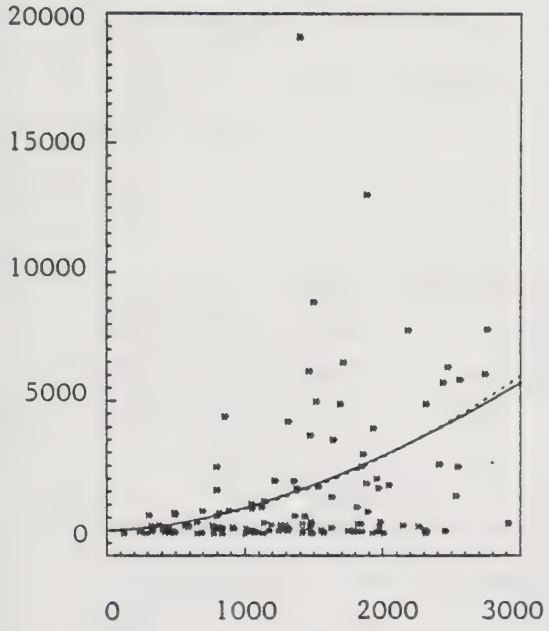


Figure 11.a.

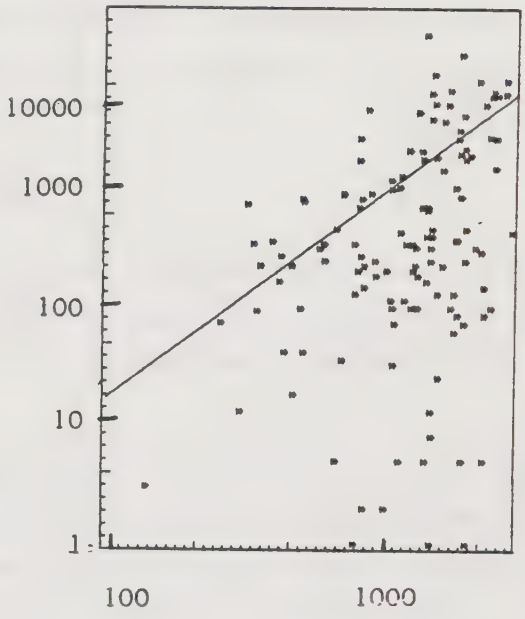


Figure 11.b.

Pour les dosages de la testostérone et du cortisol, la méthode de dosage est différente pour la préparation standard et pour les préparations inconnues. Aussi pour les analyses statistiques que nous ferons ultérieurement, il nous faut vérifier que malgré cette différence de méthode la variabilité de la variance est bien la même. Nous avons fait cette vérification pour le dosage de la testostérone à l'aide de 134 préparations inconnues ; la réponse a été mesurée deux fois pour chacune d'elles. En supposant $\phi_h(\mu) = \mu^h$, h est estimé par 1,2 avec un écart-type de 0,18, au lieu de 1,26 et 0,11 pour les préparations standards. σ^2 est estimé par 0,56 au lieu de 0,32. Sans faire de test, il semble raisonnable de supposer que la variabilité de la variance est la même pour les préparations standards et les préparations inconnues, bien que l'estimation de σ^2 soit légèrement supérieure pour les secondes. Si nous mettons en commun l'ensemble des observations, h est estimé par 1,22 avec un écart-type de 0,1 et σ^2 par 0,44.

II.5. ESTIMATION DE LA COURBE STANDARD .-

Pour chacun des exemples que nous avons déjà présentés dans ce chapitre, nous allons estimer les paramètres de la courbe standard à l'aide de deux méthodes dfifférentes : la méthode logit-log et la méthode du modèle logistique à 4 paramètres.

La première méthode est très largement utilisée et mérite donc une attention particulière, la seconde est plus satisfaisante pour le statisticien et l'occasion est donnée de montrer concrètement ses avantages. Dans le cas plus général des dosages biologiques, on pourra consulter sur cette question l'article de VOLUND (1978).

Nous nous plaçons dans le cas où nous estimons les paramètres de la courbe standard indépendamment des réponses mesurées pour les préparations inconnues, car en général le nombre de dilutions pour chacune des préparations inconnues ne permet pas de faire autrement.

II.5.1. Méthode logit-log

Soit U la réponse observée, $\bar{d}$ et $\bar{c}$ les moyennes des réponses observées aux doses nulle et infinie. Soit $U' = (U-\bar{c})/(\bar{d}-\bar{c})$. $\bar{d}$ et $\bar{c}$ étant considérées comme des constantes du modèle, la variance de U' est proportionnelle à la

à la variance de U :

$$\mathrm{Var}(U') = \mathrm{Var}(U)/(\bar{d}-\bar{c})^2$$

Soit V la variable aléatoire transformée

$$V = \mathrm{logit}(U').$$

Nous supposons que son espérance est une fonction linéaire du logarithme de la dose. $E(V) = \nu$, $\nu = \alpha + \beta x$ où $x = \mathrm{Log}(z)$. Si $\mu' = E(U')$, nous pouvons approcher la variance de V de la façon suivante :

$$\mathrm{Var}(V) = \mathrm{Var}(U')/\mu'^2\ (1-\mu')^2$$

Pour chaque dose z_i, nous observons n_i répétitions de la réponse, U_{ij} $j=1,\ldots,n_i$. Si nous supposons la variance de U proportionnelle à $\phi_h(\mu)$, nous pondérons chacun des V_{ij} par ω_i où les ω_i sont les valeurs observées pour $U_{i.}'^2\ (1-U_{i.}')^2\ (\bar{d}-\bar{c})^2/\phi_h(U_{i.}')$

avec $V_{ij} = \mathrm{logit}(U_{ij}')$, $U_{ij}' = (U_{ij}-\bar{c})/(\bar{d}-\bar{c})$

$$U_{i.}' = \sum_{j=1}^{n_i} U_{ij}'/n_i$$

Tableau 2

Tableau des données transformées. x est le logarithme de la dose, V est la réponse obtenue après la transformation logit.

$V = \text{logit}(U')$, $U' = (U-\bar{c})/(\bar{d}-\bar{c})$. ω est la variable de pondération.

Tableau 2.a. : Dosage de la prolactine $\bar{d} = 4365$, $\bar{c} = 0$.
La variance de U est constante

x	Valeurs observées pour V		$\omega x 10^{-3}$
-1,806	4,722	3,266	10
-1,505	3,360	2,782	38
-1,204	2,060	2,126	182
-0,903	0,994	1,240	660
-0,602	0,702	0,608	965
-0,523	0,208	0,376	1142
-0,398	0,076	0,020	1190
-0,301	-0,098	-0,084	1186
-0,125	-0,626	-0,588	993
0	-0,848	-0,826	848
0,097	-0,938	-0,998	758
0,398	-1,726	-1,674	325
0,699	-2,632	-2,442	89

Tableau 2.b. : Dosage de la Testostérone $\bar{d} = 1305$, $\bar{c} = 37$. La variance de U est prise proportionnelle à $\mu^{1.2}$.

x	Valeurs observées pour V		ω
0,699	2,444	2,052	3
1	1,060	1,016	16
1,301	-0,578	-0,636	50
1,699	-0,972	-2,000	33
2	-2,646	-2,712	19
2,301	-3,592	-3,592	7
2,699	-4,650	-4,834	1
3	-4,084	-4,084	3

Tableau 2.c. : Dosage du cortisol $\bar{d} = 2766$, $\bar{c} = 133$. La variance de U est prise proportionnelle à $\mu^{1.7}$.

x	Valeurs observées pour V				ω
-1,699	2,800	3,086	2,212	2,178	0,05
-1,398	2,082	2,536	1,756	2,160	0,11
-1,222	1,190	1,556	1,082	1,332	0,42
-1,097	1,112	0,998	0,920	0,946	0,63
-1	0,648	0,774	0,546	0,664	0,96
-0,699	-0,130	-0,056	-0,110	-0,222	2,01
-0,398	-0,870	-0,854	-0,974	-0,938	2,82
-0,222	-1,288	-1,290	-1,318	-1,302	2,88
-0,097	-1,572	-1,544	-1,640	-1,842	2,75
0	-1,818	-1,836	-1,892	-1,842	2,55
0,176	-2,216	-2,370	-2,186	-2,206	2,10
0,301	-2,514	-2,404	-2,446	-2,498	1,80
0,602	-3,068	-2,974	-3,124	-3,142	1,04

Figure 12

* : série des points observés après transformation logit, en fonction du logarithme de la dose (V_{ij}, x_i)

—— : droite estimée après régression pondérée par w_i de V sur x.

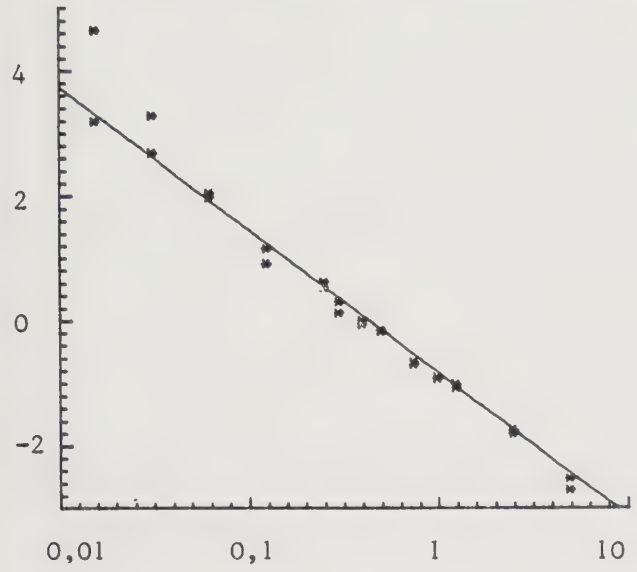


Figure 12.a.

Dosage de la prolactine. La droite estimée est d'équation $\hat{v} = -0,830-2,294x$

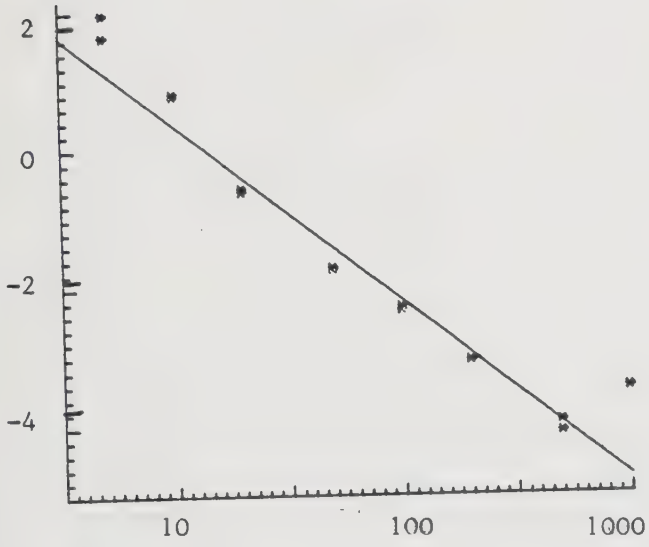


Figure 12.b.

Dosage de la testostérone. La droite estimée est d'équation $\hat{v} = 3,498-3,062x$

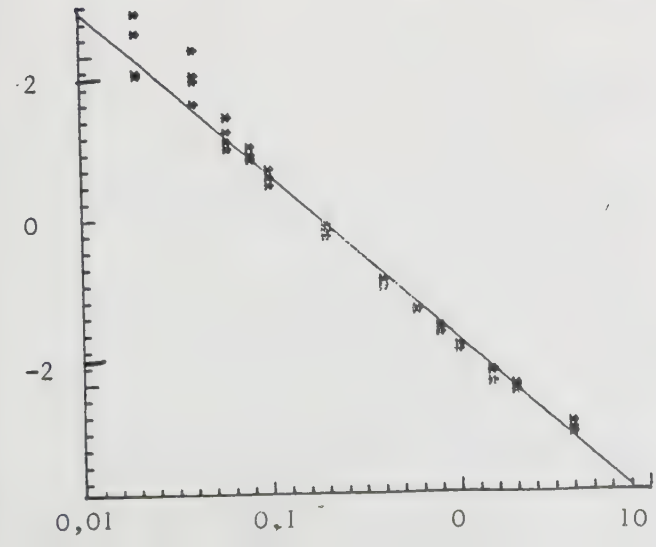


Figure 12.c.

Dosage du cortisol. La droite estimée est d'équation $\hat{v} = -1,804-2,428x$

Figure 13

Graphe des résidus réduits en fonction du logarithme de la dose. Ces graphes montrent que l'ajustement des données transformées est convenable pour la prolactine, mais n'est pas acceptable aussi bien pour la testostérone que pour le cortisol.

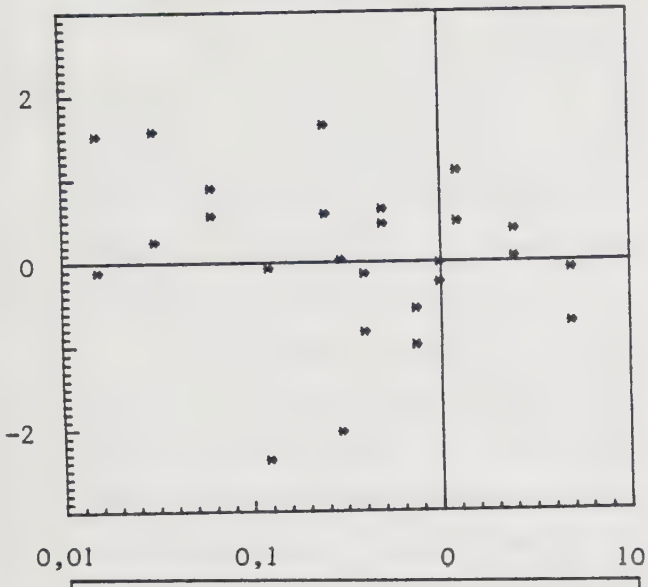


Figure 13.a.

Dosage de la prolactine. Le graphe ne semble pas présenter de structure particulière.

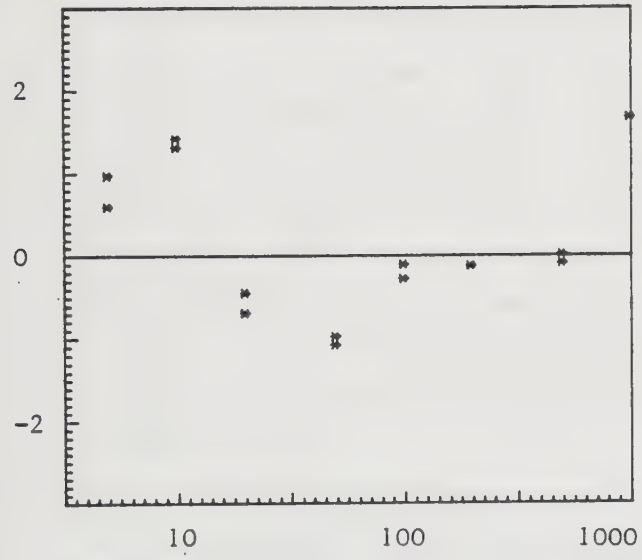


Figure 13.b.

Dosage de la testostérone. Les données sont surestimées pour les faibles doses, sous estimées ensuite.

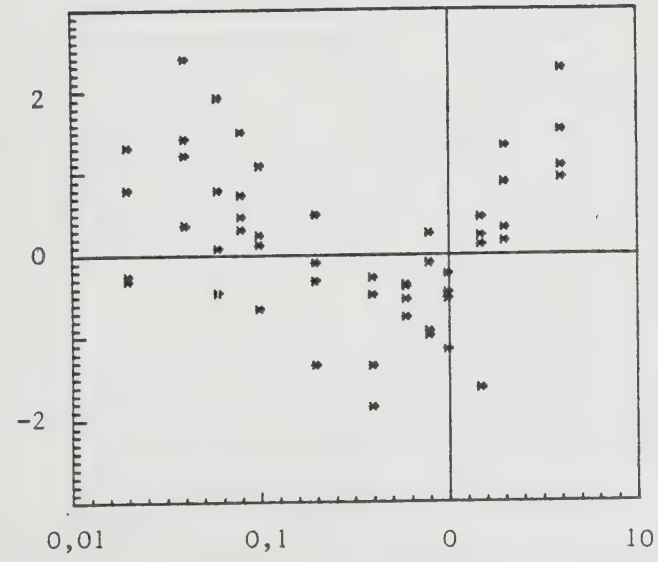


Figure 13.c.

Dosage du cortisol. A l'évidence une structure particulière existe.

La méthode logit-log est apparemment applicable au dosage de la prolactine, mais elle est inapte à décrire la courbe standard des dosages de la testostérone et du cortisol.

Ceci n'est pas une révélation. Habituellement la courbe standard du dosage de la prolactine présenté ici, est estimée à l'aide de la méthode logit-log. La pondération n'est pas prise en compte explicitement. L'expérimentateur pondère grossièrement ; il retire de l'analyse les données qui lui semblent à première vue trop éloignées d'une droite qui ajusterait au mieux les points situés au centre de la courbe.

En ce qui concerne le dosage de la testostérone, le problème est plus délicat. Le graphe des résidus (figure 13.b) montre que l'ajustement des données à une droite ne convient pas malgré la pondération utilisée. Nous pourrions penser que certaines données sont suspectes et qu'il vaut mieux les éliminer de l'analyse que de les garder, même si la valeur de la pondération allouée en ces points est très faible. A première vue, nous sommes tentés d'éliminer les réponses obtenues pour les doses 5, 10 et 1000 pg/100 µℓ. Si nous supprimons ces doses, les comptages observés pour la courbe standard varient entre 50 et 500 (les comptages sont exprimés en coups par minute). Or, pour de nombreuses préparations inconnues les comptages observés sont supérieurs à 500 ; nous garderons donc les réponses observées en 10 pg/100 µℓ. Même après élimination de ces données, l'ajustement à une droite ne convient pas (voir figure 14). Aussi pour le dosage de la testostérone, nous ne pouvons pas retenir la méthode logit-log pour estimer les paramètres de la courbe standard.

Pour le dosage du cortisol, la méthode logit-log n'est jamais appliquée, l'écart du modèle à la linéarité étant bien connu du biologiste. Celui-ci considère les données transformées U' en fonction du logarithme de la dose x (figure 4.c.) et tente d'ajuster deux droites aux observations ; le problème important de la détermination du point de rupture est alors posé. La méthode est d'autant plus suspecte que le point de rupture ne correspond à aucun phénomène biologique.

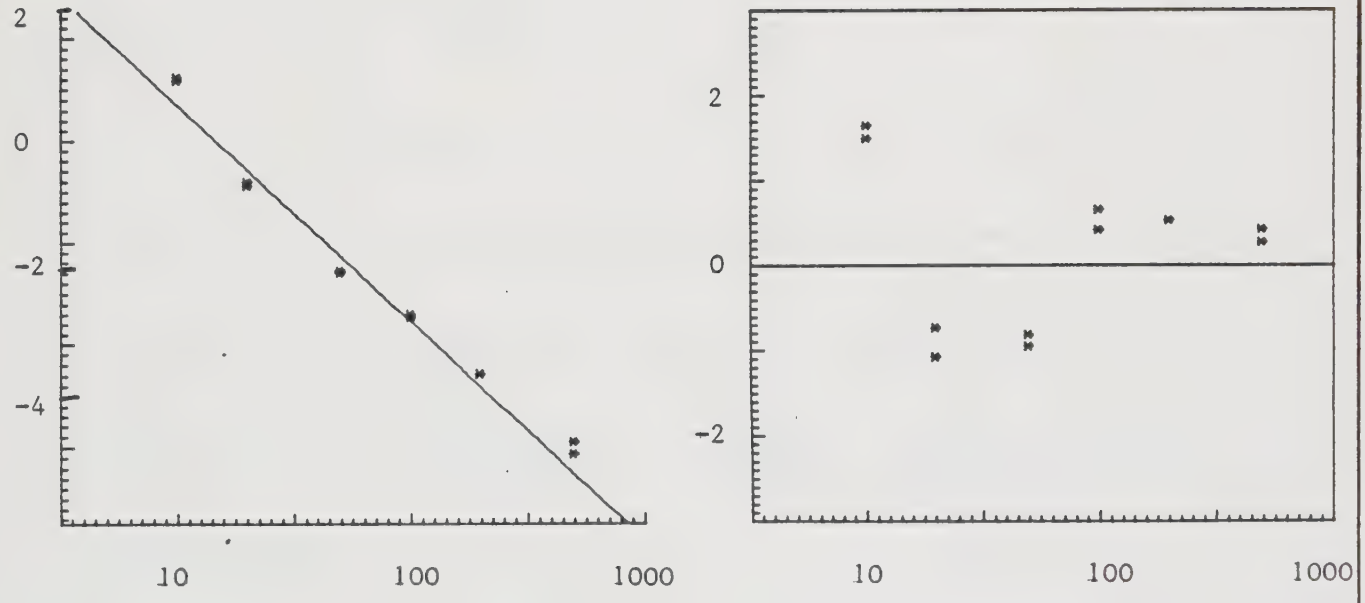
Figure 14

Estimation de la courbe standard du dosage de la testostérone après élimination des observations en 5 et 1000 pg/100 µℓ.

Figure 14.a. : ** : série des points observés après transformation logit en fonction du logarithme de la dose.

—— : droite estimée après régression pondérée par w de V sur x
$\hat{v} = 3{,}94 - 3{,}40\ x$

Figure 14.b. : Graphe des résidus réduits en fonction du logarithme de la dose.



II.5.2. Modèle logistique à quatre paramètres

Le modèle : U la variable aléatoire observée a pour espérance μ, sa variance est proportionnelle à $\phi_h(\mu)$

$$\mu = c + (d-c)\, f(y)$$

$$f(y) = 1/(1+\exp -y) \qquad y=\alpha+\beta x \qquad x=\mathrm{Log}(z)$$

$$\mathrm{Var}(U) = \sigma^2\, \phi_h(\mu) \qquad \phi_h(\mu) = \mu^h \text{ ou } \mu(1+h\mu)$$

La méthode d'estimation des paramètres α, β, c, d et σ^2 est discutée en annexe. Le lecteur y trouvera également la notice d'utilisation d'un programme conversationnel écrit en Fortran.

Bien que le modèle ne soit pas linéaire en fonction des paramètres, nous pouvons effectuer un test d'adéquation au modèle similaire au test de non-linéarité.

Soit U_{ij} la jème répétition de la réponse observée pour la dose z_i d'hormone ; soit S^2 l'estimateur de σ^2 ; soit

$$\mathcal{S}_A = \sum_{i,j} (U_{ij} - U_{i.})^2/\phi_h(U_{i.}) \quad \text{et} \quad \mathcal{S}_H = (n-4)\, S^2 \quad \text{où}$$

$n = \sum_i n_i$. $\mathcal{S}_H - \mathcal{S}_A$ traduit la variabilité due au choix du modèle (c'est-à-dire à l'ajustement des données à la logistique) et $\mathcal{S}_A$ celle due à la variabilité des mesures. Le test repose sur la comparaison de la valeur calculée pour la statistique $\mathcal{U} = \dfrac{(\mathcal{S}_H - \mathcal{S}_A)/(k-4)}{\mathcal{S}_A/(n-k)}$ à la valeur d'un Fisher à (k-4) et (n-k) degrés de liberté.

1er exemple : Dosage de la prolactine

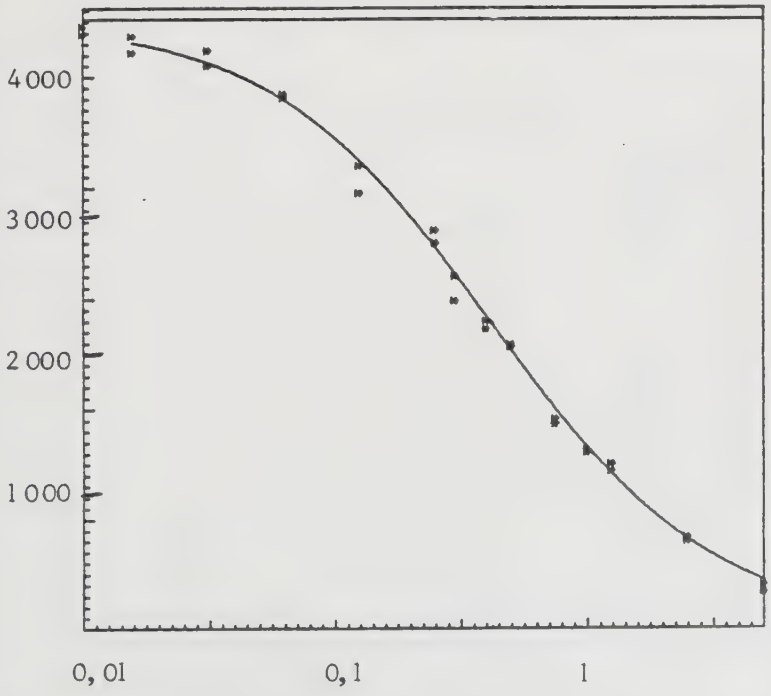
Conformément à l'étude faite au paragraphe II.4., nous avons considéré la variance de la réponse constante et égale à σ^2.

	paramètres estimés	écarts-types estimés
c	1,7	102
d	4418	47
α	-0,8492	0,0862
β	-2,2680	0,1172

σ^2 a été estimé par 7039,6 avec 24 degrés de liberté. La valeur calculée de la statistique $\mathcal{U}$ est 2,4 ; la valeur d'un Fisher au niveau 5 % de degrés de liberté 10 et 14 vaut 2,6 ; nous sommes donc satisfait de l'ajustement de la courbe standard (voir figure 15).

Nous avions supposé pour ce dosage que la valeur de c était nulle. Notre hypothèse se trouve confortée par nos résultats.

Figure 15



Dosage de la prolactine

* : réponses observées en fonction du logarithme de la dose

—— : courbe et asymptotes estimées

Toutefois, remarquons que l'écart-type estimé pour c est très grand par rapport à l'estimation. Ceci est dû au plan d'expérience qui ne permet pas l'estimation de ce paramètre (voir paragraphe II.7.). Aussi nous avons intérêt pour ce dosage à supposer c nul et à estimer seulement les trois paramètres d, α, β. Les résultats sont alors les suivants

	paramètres estimés	écarts-types estimés
c=0		
d	4419	42
α	-0,8482	0,02846
β	-2,2660	0,0640

σ^2 a été estimé par 6578,1 avec 25 degrés de liberté. La statistique $\mathcal{U}$ vaut 2,2.

2ème exemple : Dosage de la testostérone

Nous avons supposé la variance proportionnelle à $\mu^{1,2}$

	paramètres estimés	écarts-types estimés
c	48	3,6
d	1331	33
α	4,5336	0,308
β	-3,8430	0,2052

σ^2 a été estimé par 0,88 avec 19 degrés de liberté.
La statistique $\mathcal{U}$ vaut 36. La valeur d'un Fisher au niveau 5 % de degrés de liberté 6 et 13 vaut 3 ; nous ne pouvons nous satisfaire de l'ajustement effectué. La figure 16.a. montre que la réponse pour les faibles doses est sous-estimée.

Figure 16 : Dosage de la testostérone

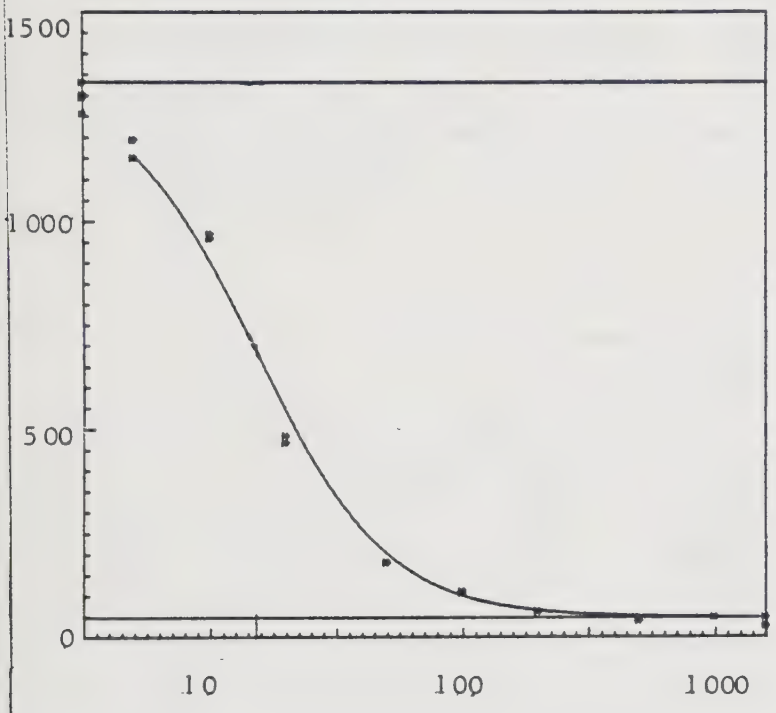


Figure 16.a.

* : réponses observées en fonction du logarithme de la dose

—— : courbe et asymptotes estimées

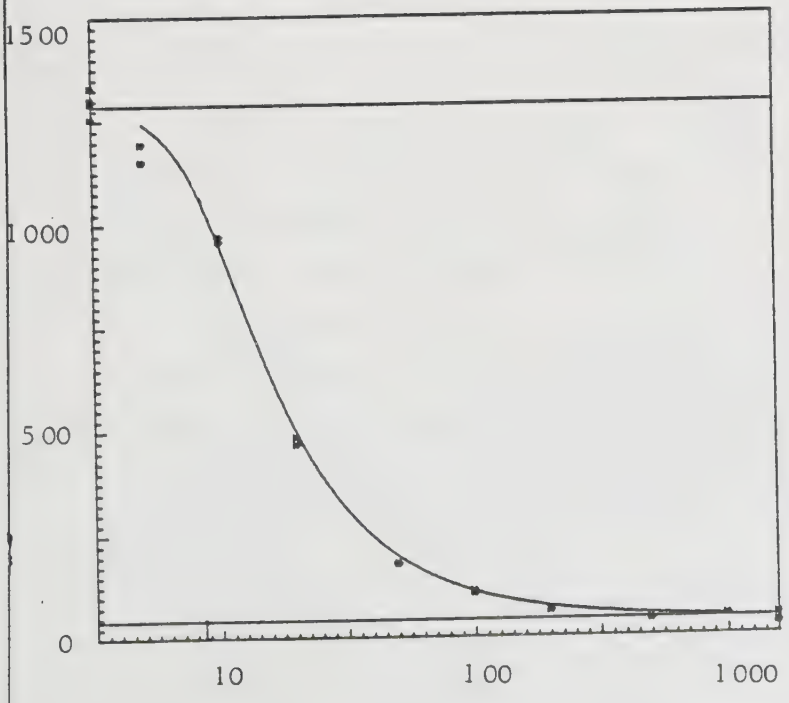


Figure 16.b. γ=.3

* : réponses observées en fonction du logarithme de la dose

——: courbe et asymptotes estimées

Pour traduire la dissymétrie de la courbe nous pouvons essayer de modifier le modèle de la façon suivante :

$\mu = c + (d-c)\ f(y)^{\gamma}$ si γ est positif ou bien

$\mu = c + (d-c)\ (1-(1-f(y))^{-\gamma}$ si γ est négatif. Ainsi la courbe de réponse s'approchera plus vite d'une des deux asymptotes, plus lentement de l'autre ; cela dépendra de la valeur de γ. Comme nous l'avons fait pour h, nous allons considérer γ comme un paramètre de second niveau, pour les mêmes raisons que celles que nous avons déjà exposées au paragraphe II.4. ; souvent l'utilisateur n'a aucune idée a priori de la valeur de γ, un programme peut lui en fournir une estimation (voir annexe C). Ainsi pour le dosage de la testostérone, une valeur de γ égale à 0,3 fournit les résultats suivants.

$\gamma = 0,3$	paramètres estimés	écarts-types estimés
c	42	2,3
d	1285	20
α	8,3612	0,4122
β	-8,9956	0,296

σ^2 a été estimé par 0,40 , la statistique $\mathcal{U}$ vaut 14,5.

Nous n'avons pu améliorer ce résultat, en essayant l'ajustement avec d'autres valeurs de γ, ce qui signifie, bien que la valeur de la statistique $\mathcal{U}$ ait fortement diminué, que le modèle n'est pas bon (figure 16.b.). Pratiquement, que devons nous faire ? Nous avons plusieurs choix possibles :

- stopper l'analyse, c'est-à-dire "jeter le dosage". En général, un simple examen des données permet au biologiste de détecter s'il y a eu une anomalie technique lors du dosage (il peut s'agir par exemple d'une erreur de dilution de l'anticorps, ou d'une quantité d'hormone radioactive H^* trop faible). Si les données semblent présenter une telle anomalie, sans hésitation, le dosage est rejeté.

- Pour les données du dosage de la testostérone traitées ici, il n'y a aucune raison de faire ce choix. Examinons une nouvelle fois la figure 16.b. On constate que les observations de la réponse pour la dose z=5 s'écartent de la courbe ajustée.

Nous pouvons alors formuler deux suppositions :

1. La modélisation de la courbe standard est bonne, mais les deux observations de la réponse pour la dose z=5 sont erronées. Si cela est vrai, il nous faut recommencer l'analyse en supprimant ces observations qui sont alors considérées comme des valeurs aberrantes. Dans notre cas précis, nous n'avons aucune raison de supposer ces données aberrantes.

2. Ces deux observations ne sont pas erronées, la modélisation de la courbe standard que nous avons choisie est donc inapte à décrire le phénomène biologique pour les faibles doses. Nous pourrions alors chercher une autre modélisation de la courbe standard. Une première possibilité consiste à chercher une autre fonction de régression de la logistique. Mais si ceci nous conduit à augmenter le nombre de paramètres, nous diminuerons encore la puissance de notre test qui est déjà faible (6 et 13 degrés de liberté pour la statistique $\mathcal{U}$). Une seconde possibilité consiste à utiliser une méthode non paramétrique telle que la méthode des splines par exemple. Nous pensons que nous ne pouvons attendre une amélioration très sensible de l'ajustement à l'aide d'une telle étude. Aussi, nous garderons ce modèle pour estimer dans le paragraphe suivant les doses inconnues en évitant toutefois de tenir compte des résultats qui auront été estimés à l'aide de la partie de la courbe standard proche de l'asymptote supérieure.

De plus, nous constatons que pour ce dosage, la variabilité due aux mesures est très faible ; il est donc vivement conseillé de modifier le plan d'expérience en augmentant le nombre de répétitions en chaque dose, afin d'obtenir une meilleure estimation de cette variabilité.

3ème exemple : Dosage du cortisol

Nous avons supposé la variance proportionnelle à $\mu^{1,7}$.

	paramètres estimés	écarts-types estimés
c	140	3,6
d	2836	30
α	-1,859	0,0224
β	-2,4316	0,0374

σ^2 a été estimé par 0,0135 avec 60 degrés de liberté. La statistique $\mathcal{U}$ vaut 4,86. La valeur d'un Fisher au niveau 5 % de degrés de liberté 11 et 49 vaut 2. Comme pour la testostérone, la courbe standard présente une dissymétrie que nous allons tenter de traduire en introduisant le paramètre γ (sur la figure 17.a. nous remarquons que les réponses pour les fortes doses ont tendance à être surestimées alors que les réponses pour les faibles doses sont sous estimées). En posant γ égal à 0,6 nous obtenons les résultats suivants :

	paramètres estimés	écarts-types estimés
c	133	2,7
d	2758	21,4
α	-3,2466	0,0242
β	-3,2934	0,0368

σ^2 a été estimé par 0,0077, $\mathcal{U}$ vaut 0,95, le modèle semble satisfaisant.

Figure 17

Dosage du Cortisol

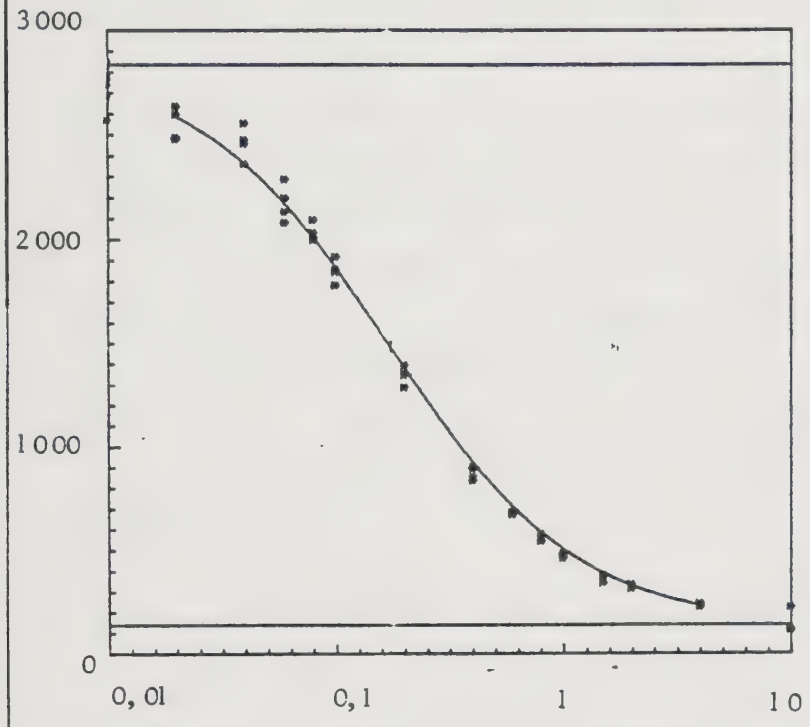


Figure 17.a.

* : réponses observées en fonction du logarithme de la dose

—— : courbe et asymptotes estimées

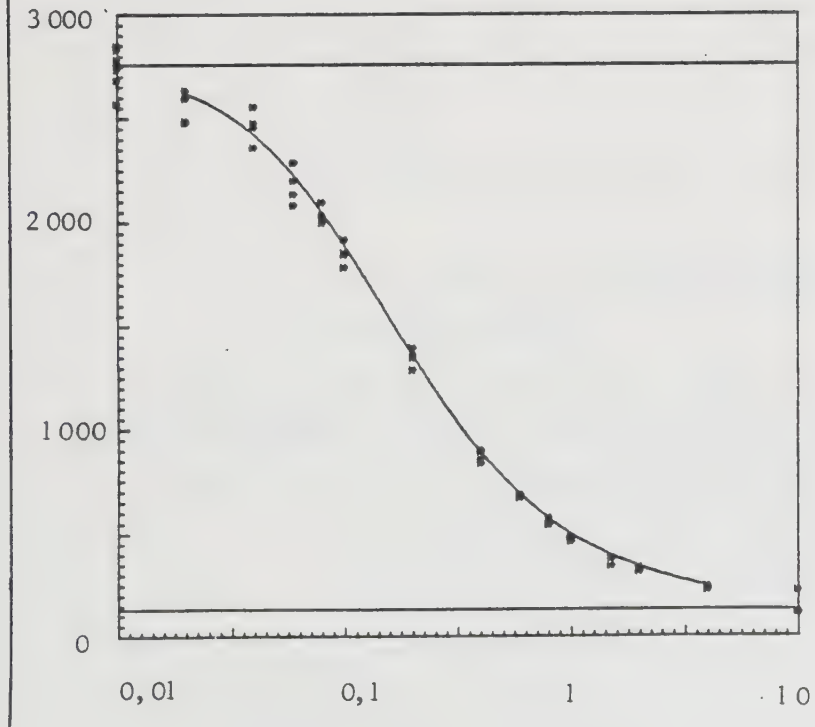


Figure 17.b. γ=.6

* : réponses observées en fonction du logarithme de la dose

——: courbe et asymptotes estimées

II.6. ESTIMATION DES DOSES INCONNUES ET CALCUL D'UN INTERVALLE DE CONFIANCE .-

II.6.1. Estimation

Supposons encore une fois que la réponse pour la préparation inconnue étudiée soit mesurée en une seule dilution ($\ell=1$).

Supposons de plus, pour simplifier les notations que cette réponse soit mesurée pour un volume de préparation inconnue égal au volume de dosage de la préparation standard (il suffira, une fois les estimations calculées de corriger par un facteur multiplicatif).

Soit U^I la réponse pour une préparation où la dose inconnue d'hormone est z^I, soit $x^I=\text{Log } z^I$. Si nous supposons les conditions de validité du dosage vérifiées et si nous notons $\mu^I=E(U^I/z^I)$

alors $\mu^I = c + (d-c)f(y^I)$ où $y^I = \alpha + \beta x^I$

En inversant la relation entre μ^I et x^I, nous pouvons exprimer x^I de la façon suivante :

$$x^I = \frac{y^I-\alpha}{\beta} \quad \text{où } y^I = -\ell n((d-\mu^I)/(\mu^I-c))$$

Si $\overset{\cdot}{U}{}^I$ est la moyenne des r^I réponses pour la préparation inconnue I, les paramètres α,β,d et c ayant été estimés par A,B,D et C à l'aide des réponses observées pour la préparation standard, l'estimateur de x^I est $X^I = (Y^I-A)/B$ où $Y^I = -\ell n((D-\overset{\cdot}{U}{}^I)/(\overset{\cdot}{U}{}^I-C))$

Nous allons étudier deux méthodes de calcul d'un intervalle de confiance pour x^I.

II.6.2. 1ère méthode : calcul d'un intervalle de confiance à l'aide du théorème de Fieller

Pour calculer un intervalle de confiance pour x^I, nous allons, comme le propose FINNEY (1978) appliquer le théorème de Fieller au rapport $(y^I-\alpha)/\beta$.

Nous devons faire les hypothèses suivantes :

. les estimateurs Y^I-A et B sont normalement distribués

. ce sont des estimateurs sans biais de $y^I-\alpha$ et β.

$$E(Y^I-A) = y^I-\alpha \qquad E(B) = \beta$$

. ils doivent vérifier

$$\mathrm{Var}(Y^I - A) = \sigma^2 v_{11}$$

$$\mathrm{Var}(B) = \sigma^2 v_{22}$$

$$\mathrm{Cov}(Y^I - A, B) = \sigma^2 v_{12}$$

où les v_{ij} ne dépendent que du plan expérimental (choix des doses de la préparation standard et nombre de répétitions de la réponse observée en chacune de ces doses, nombre de répétitions r des observations de la réponse pour la préparation inconnue I). σ^2 a été défini par $\mathrm{Var}(U) = \sigma^2 \phi_h(\mu)$, ce qui signifie que la variabilité de la variance est la même pour la préparation standard et pour la préparation inconnue I. Nous reviendrons sur la validité de ces hypothèses, supposons les vérifiées, et calculons les v_{ij}.

$$\mathrm{Var}(Y^I - A) = \mathrm{Var}(Y^I) + \mathrm{Var}(A) - 2\,\mathrm{Cov}(Y^I, A)$$

Y^I est une fonction de trois variables aléatoires C,D, $U_\cdot^I$ que nous allons noter $g(C,D,U_\cdot^I)$. Nous allons chercher un estimateur de la variance de Y^I en développant la fonction g au voisinnage du point (c,d,U^I).

$$\mathrm{Var}(Y^I) = \left(\frac{\partial g}{\partial U^I}\right)^2 \mathrm{Var}(U_\cdot^I) + \left(\frac{\partial g}{\partial c}\right)^2 \mathrm{Var}(C) + \left(\frac{\partial g}{\partial d}\right)^2 \mathrm{Var}(D)$$

$$+ 2\,\frac{\partial g}{\partial U^I}\,\frac{\partial g}{\partial c}\,\mathrm{Cov}(U_\cdot^I, C) + 2\,\frac{\partial g}{\partial U^I}\,\frac{\partial g}{\partial d}\,\mathrm{Cov}(U_\cdot^I, D) + 2\,\frac{\partial g}{\partial c}\,\frac{\partial g}{\partial d}\,\mathrm{Cov}(C,D)$$

les dérivées de g étant calculées au point $(C,D,U_\cdot^I)$.

Les réponses étant mesurées indépendamment les unes des autres, C et D dépendant seulement des réponses pour la préparation standard ,

$$\mathrm{Cov}(U_\cdot^I, C) = \mathrm{Cov}(U_\cdot^I, D) = 0$$

Finalement, un estimateur de la variance de Y^I sera

$$\mathrm{Var}(Y^I) = \frac{(D-C)^2}{(U_\cdot^I - C)^2 (D - U_\cdot^I)^2}\,\mathrm{Var}(U_\cdot^I) + \frac{1}{(U_\cdot^I - C)^2}\,\mathrm{Var}(C)$$

$$+ \frac{1}{(D-U_\cdot^I)^2}\,\mathrm{Var}(D) + \frac{2\,\mathrm{Cov}(C,D)}{(U_\cdot^I - C)(D - U_\cdot^I)}$$

Nous avons supposé que la variabilité des réponses est la même pour la préparation standard et pour les préparations inconnues. Cette hypothèse implique que $Var(U_{\cdot}^{I}) = \sigma^2 \phi_h(\mu^I)/r^I$ que nous estimons par $S^2 \phi_h(U_{\cdot}^{I})/r^I$.

De la même manière, les estimateurs de $Cov(Y^I,A)$ et $Cov(Y^I,B)$ sont

$$Cov(Y^I,A) = \frac{-1}{U_{\cdot}^{I}-C} Cov(A,C) - \frac{1}{D-U_{\cdot}^{I}} Cov(A,D)$$

$$Cov(Y^I-A,B) = \frac{-1}{U_{\cdot}^{I}-C} Cov(B,C) - \frac{1}{D-U_{\cdot}^{I}} Cov(B,D) - 2Cov(A,B)$$

Nous pouvons maintenant appliquer le théorème de Fieller (Annexe du Chapitre 1). Si nous notons X_{inf} et X_{sup} les bornes de l'intervalle de confiance au niveau α pour x^I,

$$X_{inf}, X_{sup} = \left[X^I - H\, v_{12}/v_{22} \pm \mathcal{T}\, \Sigma \sqrt{G}/B\right] /(1-H)$$

où $H = \mathcal{T}^2\ \Sigma^2\, v_{22}/B^2$

et $G = v_{11} - 2\, X^I\, v_{12} + (X^I)^2\, v_{22} - H(v_{11}-v_{12}^2/v_{22})$

$\mathcal{T} = \mathcal{T}(\alpha, n-4)$

z^I est estimé par 10^{X^I} et avec une probabilité $1-\alpha$, appartient à l'intervalle $[10^{X_{inf}}; 10^{X_{sup}}]$.

1er exemple : Dosage de la prolactine

Soit v^I le volume pour lequel les r^I réponses de la préparation inconnue I ont été mesurées.

tableau des résultats

r^I	v^I en µl	$U^I_.$	z^I en ng/5µl	IC à 95 %(*)	z^I en ng/ml	IC à 95 %(*)
2	5	4300	0,011	[0,003 ; 0,037]	2,2	[0,6 ; 7,4]
4	5	4300	0,011	[0,004 ; 0,029]	2,2	[0,8 ; 5,8]
2	200	2200	0,011	[0,009 ; 0,012]	2,2	[1,8 ; 2,4]
2	5	2200	0,426	[0,378 ; 0,479]	85,2	[75,6 ; 95,8]
4	5	2200	0,426	[0,390 ; 0,465]	85,2	[78,0 ; 93,0]
2	3	300	6,05	[3,85 ; 9,64]	1210	[770 ; 1928]
4	5	300	6,05	[4,34 ; 8,56]	1210	[868 ; 1712]
2	0,07	2200	6,09	[5,40 ; 6,85]	1218	[1080 ; 1370]

La longueur de l'intervalle de confiance est réduite d'environ 25 % en effectuant 4 répétitions au lieu de 2. L'examen du tableau montre également combien il est important que la réponse moyenne observée soit proche du milieu des deux asymptotes : pour une faible dose de 2,2 ng/ml, diluer la préparation afin que la réponse observée se situe à proximité du milieu des deux asymptotes diminue la longueur de l'intervalle de confiance de 90 % ; pour une forte dose de l'ordre de 1200 ng/ml, le gain est de 75 %.

(*) *Intervalle de confiance calculé au niveau 95 %.*

2ème exemple : Dosage de la testostérone

L'introduction du paramètre γ rend les calculs des v_{ij} plus laborieux ; et il n'est pas intéressant de les détailler ici.

Tableau des résultats

r^I	v^I en ml	valeur observée de $U^I_.$	valeur calculée pour z^I en pg/ml	Intervalle de confiance
2	0,1	1200	60,3	[460 ; 783]
4	0,1	1200	60,3	[487 ; 739]
2	0,25	650	61,2	[559 ; 669]
2	0,1	650	153	[140 ; 167]
4	0,1	650	153	[142 ; 165]
2	0,1	70	2170	[1480 ; 3230]
4	0,1	70	2170	[1620 ; 2950]
2	0,007	650	2190	[2000 ; 2390]

La longueur de l'intervalle de confiance est réduite d'environ 20 % (avec des inégalités selon la valeur des réponses moyennes) en effectuant 4 répétitions au lieu de 2. Pour une dose de 60 pg/ml, diluer la préparation afin que la réponse observée se situe environ au milieu des deux asymptotes permet un gain de 60 % sur la longueur de l'intervalle de confiance ; pour une dose de l'ordre de 2200 pg/ml le gain est de 78 %. Pour toute réponse moyenne inférieure à 50 (ou bien pour toute concentration d'hormone supérieure à 6000 pg/ml), les valeurs pour X_{inf} ou X_{sup} sont trop grandes pour que les limites calculées pour z^I soient représentables en machine ; nous ne pouvons donc donner aucun résultat pour ces préparations inconnues.

3ème exemple : Dosage du cortisol

r^I	v^I en ml	valeur observée pour $U_\bullet^I$	z^I valeur calculée pour x^I en ng/ml	intervalle de confiance
2	0,1	2600	0,22	[0,13 ; 0,46]
4	0,1	2600	0,22	[0,15 ; 0,32]
2	0,7	1600	0,21	[0,19 ; 0,23]
2	0,1	1600	1,46	[1,33 ; 1,59]
4	0,1	1600	1,46	[1,36 ; 1,56]
2	0,1	250	38,8	[33,5 ; 44,8]
4	0,1	250	38,8	[34,8 ; 43,2]
2	0,0038	1600	38,4	[35,0 ; 41,8]

La longueur de l'intervalle de confiance est réduite d'environ 50 % pour les faibles doses, 40 % pour les fortes doses et 20 % pour les doses moyennes en effectuant 4 répétitions au lieu de 2. Pour une dose d'environ 0,2 ng/ml, diluer la préparation afin que la réponse observée se situe environ au milieu des 2 asymptotes permet un gain de 90 % sur la longueur de l'intervalle de confiance ; pour une dose de l'ordre de 40 ng/ml le gain est de 40 %.

Certes nous avons calculé un intervalle de confiance pour des quantités inconnues d'hormone grâce au théorème de Fieller. Mais que vaut ce calcul ? Nous ferons trois remarques :

1) Si les conditions d'application du théorème de Fieller étaient vérifiées, la probabilité de se tromper en affirmant que x^I appartient à l'intervalle X_{inf}, X_{sup}

serait exactement égale à α. Malheureusement nous ne savons pas si les conditions du théorème de Fieller sont vérifiées. Pire, nous avons de bonnes raisons d'en douter, et nous ne savons pas dans quelle mesure nous nous en écartons. Le problème est théoriquement délicat pour deux raisons : l'espérance de la réponse n'est pas une fonction linéaire des paramètres, et la variance des réponses n'est pas constante et dépend de l'espérance. Par exemple, le théorème de Fieller suppose que les estimateurs Y^I-A et B de y^I-α et β sont non biaisés. Or quelle que soit la méthode d'estimation utilisée (voir annexe B, estimation par moindres carrés ou par maximum de vraisemblance), le nombre d'observations étant fini (de l'ordre de quelques dizaines), ces estimateurs sont sûrement biaisés. Que vaut ce biais ? Quelles sont les conséquences de ce biais sur le calcul de l'intervalle de confiance à l'aide du théorème de Fieller ? Tant que nous n'avons pas répondu à ces questions, nous ne pouvons affirmer que le théorème de Fieller soit une bonne méthode pour calculer les intervalles de confiance autour des doses inconnues. Une autre façon de tester la validité du théorème de Fieller serait d'effectuer de nombreuses simulations de courbes standard et de doses inconnues pour un modèle donné. Cette recherche est en cours.

2) Pourvu que H soit inférieur à 1 (si jamais ce n'était pas le cas, il faudrait jeter ce dosage et recommencer), l'intervalle de confiance calculé est toujours fini, seule la capacité du calculateur nous impose des limites. Or pour les dosages radioimmunologiques nous savons que pour les faibles doses, par exemple, il existe une dose limite en deça de laquelle la courbe "ne décroche pas", ce qui signifie que la méthode de dosage utilisée ne permet pas de doser des quantités d'hormone inférieures à cette limite. Cela n'a aucun sens de calculer un intervalle de confiance borné aux deux extrémités pour ces doses.

3) Le théorème de Fieller suppose également que la variabilité de la variance est la même pour la préparation standard et pour les préparations inconnues. Or nous avons parfois de bonnes raisons, basées sur des considérations biologiques et non statistiques de mettre en cause ce phénomène (voir paragraphe II.8.). Le paragraphe qui suit propose une autre méthode de calcul d'intervalle de confiance qui tente de prendre en compte l'ensemble de ces critiques.

II.6.3. 2ème méthode : Méthode de calcul d'un intervalle de confiance proposée par ROSS (1978).

Rappelons le modèle qui nous permet d'estimer la courbe standard :

$$U = \mu + E$$

μ l'espérance de U s'écrit $F(x,\theta)$ où θ est le vecteur des paramètres c,d,α,β.

E l'erreur résiduelle est supposée Normalement distribuée d'espérance nulle et de variance $\sigma^2 \phi_h(\mu)$.

Soit maintenant U^I la réponse pour la préparation inconnue I et x^I la dose que nous cherchons à estimer $U^I = \mu^I + E^I$ où $\mu^I = F(x^I,\theta)$, E^I étant Normalement distribuée, d'espérance nulle et de variance $\tau^2 \phi_h(\mu^I)$. x^I est une fonction des paramètres θ et μ^I. Plus précisément, $x^I = F^{-1}(\mu^I,\theta)$. On choisira comme estimateur de x^I, $X^I = F^{-1}(U^I_\cdot, \Theta)$ où $U^I_\cdot$ est la moyenne des r^I réponses pour la préparation I ($U^I_\cdot$ est un estimateur sans biais de μ^I) et Θ est le vecteur des estimateurs C,D,A, et B de c,d,α et β. Notons ψ le vecteur formé des paramètres μ^I et θ :

$$x^I = F^{-1}(\psi)$$

La méthode de calcul d'un intervalle de confiance pour x^I se décompose en deux étapes :

1) Chercher une région de confiance pour ψ au niveau $1-\alpha$ (Nous la noterons $\mathcal{R}_\alpha$).

2) En déduire un intervalle de confiance pour x^I en cherchant les extrêmum de F^{-1} pour les valeurs de ψ appartenant à $\mathcal{R}_\alpha$.

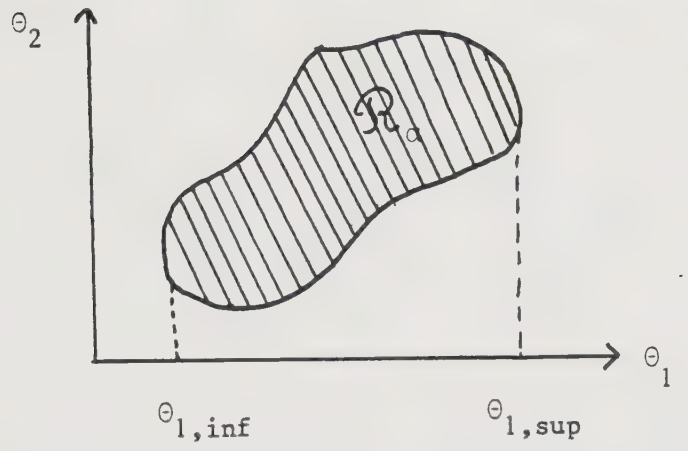
Déjà, nous devons faire une première remarque : l'intervalle de confiance que nous déterminerons ainsi pour x^I sera un intervalle de confiance conservatif, c'est-à-dire de niveau supérieur ou égal à $1-\alpha$.

Pour comprendre facilement cela, supposons que $\mathcal{R}_\alpha$ soit une région de confiance de niveau $1-\alpha$ pour le couple de praramètres (θ_1,θ_2).

L'intervalle $[\Theta_{1,inf}, \Theta_{1,sup}]$ est un intervalle de confiance pour θ_1 obtenu en projetant $\mathcal{R}_\alpha$ sur l'axe Θ_1.

$$P(\theta_1 \in [\Theta_{1,inf}, \Theta_{1,sup}] = P(\theta_1 \in [\Theta_{1,inf}, \Theta_{1,sup}] \text{ et } \theta_2 \in \mathbf{R})$$

$$\geq P((\theta_1,\theta_2) \in \mathcal{R}_\alpha) = 1-\alpha.$$



Recherche d'une région de confiance pour ψ

Appelons Ψ un estimateur de ψ.

Définir une région de confiance pour ψ, c'est chercher l'ensemble des valeurs de ψ non significativement différentes de la valeur calculée pour Ψ. ψ est estimé à l'aide des n observations de la réponse pour la préparation standard et des r^I observations pour la préparation inconnue I. Soit $L(\psi)$ la vraisemblance de l'échantillon des $n+r^I$ observations par rapport à ψ et supposons que Ψ soit l'estimateur du maximum de vraisemblance de ψ. On peut démontrer que sous l'hypothèse "$\psi=\psi_o$" , la statistique $-2\ell n \dfrac{L(\psi_o)}{L(\Psi)}$

suit la loi d'un χ^2 (dont le nombre de degrés de liberté est égal à la dimension de ψ) pourvu que la taille de l'échantillon soit suffisamment grande. Ainsi, la région limitée par le contour le long duquel la vraisemblance vérifie $-2\ell n \dfrac{L(\psi)}{L(\Psi)} = \chi^2_\alpha$ (5) est une région de confiance pour ψ, dont le niveau vaut $1-\alpha$ pourvu que $n+r^I$ soit grand. Il convient donc de calculer $L(\psi)$. Les n observations de la réponse pour la préparation standard sont indépendantes des r^I observations pour la préparation inconnue I ; par conséquent $L(\psi)$ vaut $L_1(\mu^I)\ L_2(\theta)$ où $L_1(\mu^I)$ (respectivement $L_2(\theta)$) est la vraisemblance de l'échantillon des r^I (respectivement n) observations par rapport à μ^I (respectivement θ). Ainsi, maximiser L en ψ revient à maximiser L_1 en μ^I et L_2 en θ. On peut montrer que $U.^I$ est l'estimateur du maximum de vraisemblance de $\mu^{I(*)}$ et nous supposerons que Θ est l'estimateur

---

(*) $L_1(\mu) = (2\pi\tau^2\ \mu^h)^{-r/2}\ \exp - \sum_{j=1}^{r} (U_j-\mu)^2 \ /\ 2\ \tau^2\ \mu^h$. *On trouve que les valeurs de* μ *et* τ^2, *solutions du système d'équations* $\begin{cases} \dfrac{\partial L_1}{\partial \mu} = 0 \\ \dfrac{\partial L_1}{\partial \tau^2} = 0 \end{cases}$ *sont* $U.$ *et*

$\dfrac{1}{r}\ \dfrac{\Sigma(U_j-U.)^2}{U.^h}$.

du maximum de vraisemblance de θ. Donc $\Psi = (U^I_\cdot, \Theta)$ et $X^I = F^{-1}(\Psi)$.

De plus nous supposons $n+r^I$ grand. Nous avons ainsi déterminé $\mathcal{R}_\alpha$ comme étant l'ensemble des valeurs de ψ vérifiant $-2\ell n \frac{L(\psi)}{L(\Psi)} \leq \chi^2_\alpha$ (5).

ou bien l'ensemble des valeurs de μ^I et θ vérifiant

$$-2\ \ell n \frac{L_1(\mu^I)}{L_1(U^I_\cdot)} - 2\ \ell n \frac{L_2(\theta)}{L_2(\Theta)} \leq \chi^2_\alpha \ (5).$$

Il reste à chercher les extrêmum de F^{-1} pour les valeurs de μ^I et θ appartenant à $\mathcal{R}_\alpha$. Malheureusement, cette recherche est très fastidieuse (le calcul est long et coûteux en temps machine). Une telle méthode est complètement inadaptée pour des dosages avec 200, voire 500 préparations inconnues. Nous allons donc chercher à simplifier ces calculs.

Recherche d'une région de confiance approchée pour ψ

Un développement limité au second ordre de $\ell n\ L(\psi)$ au voisinnage de Ψ nous conduit à approcher la région de confiance pour ψ par l'ensemble des ψ vérifiant $(\psi-\Psi)' \frac{\partial^2 \ell n\ L}{\partial\ \psi^2}(\Psi)\ (\psi-\Psi) \leq \chi^2_\alpha$ (5).

Appelons $\mathcal{R}_{1\alpha'}$ cette nouvelle région de niveau $1-\alpha'$. α' est plus ou moins proche de α , suivant que l'approximation faite pour $\ell n\ L(\psi)$ est bonne ou non.

Pour la préparation inconnue, nous supposerons que

$-2\ \ell n \frac{L_1(\mu^I)}{L_1(U^I_\cdot)}$ vaut $\frac{r^I}{T^2\ \phi_h(U^I_\cdot)}\ (\mu^I - U^I_\cdot)^2$.

Si dans l'espace des paramètres θ, la fonction $F(x,\theta)$ au voisinnage de Θ peut être supposée linéaire alors

$-2\ \ell n \frac{L_2(\theta)}{L_2(\Theta)}$ est proche de $(\theta-\Theta)' \frac{\partial^2 \ell n\ L_2(\Theta)}{\partial\ \theta^2}(\theta-\Theta)$ et α' est proche de α.

Les mesures de non-linéarité de BEALE (1960) ou celles de BATES (1980) donnent des indications sur la validité d'une telle approximation. Le calcul de ces mesures nous autorise, pour les trois exemples étudiés, à supposer α proche de α'. Nous supposerons α égal à α'.

Recherche d'une région de confiance pour x^I.

Afin de déterminer une région de confiance pour x^I, nous devons maintenant calculer les extrêma de F^{-1} sur le contour d'une région de confiance pour ψ. Or, le calcul des extrêma de F^{-1} sur le contour de $\mathcal{R}_{1\alpha}$ est très simple à programmer (voir annexe D).

remarques

1) Comme nous le montre le calcul, les bornes de l'intervalle de confiance pour x^I peuvent être infinies pour des doses trop faibles ou trop fortes.

Si $X_{inf} = -\infty$, la borne inférieure pour z^I vaut zéro ; la dose n'est pas significativement différente de zéro.

Si $X_{sup} = +\infty$, la borne supérieure pour z^I vaut l'infinie : la dose n'est pas significativement différente de la dose infinie.

2) Nous ne supposons pas que $Y^I = -\ell n((D-U^I_\cdot)/(U^I_\cdot-C))$ est un estimateur sans biais de y^I. Nous ne faisons aucune hypothèse particulière sur les variances et covariances de Y^I-A et B. Nous n'exigeons pas que la variabilité des préparations inconnues soit identique à celle de la préparation standard, et ceci est avantageux par rapport à la méthode du théorème de FIELLER.

3) Toutefois, pour que la convergence en loi du rapport des logarithmes de la vraisemblance vers un χ^2 soit vraie, l'estimateur du maximum de vraisemblance Ψ doit être asymptotiquement Normalement distribué et efficace. Donc nous n'exigeons pas que Θ soit un estimateur sans biais de θ. Il suffit que la loi de $\sqrt{n}\ (\Theta-\theta)$ tende, quand n devient grand, vers la loi d'une variable aléatoire

Normale d'espérance nulle et de variance minimum (DACUNHA-CASTELLE et DUFLO, 1982, p. 173).

4) Nous ne connaissons pas le niveau exact de l'intervalle de confiance calculé pour x^I, ceci pour trois raisons :

* la région de confiance pour ψ, notée $\mathcal{R}_\alpha$, est de niveau $1-\alpha$, seulement si le nombre des observations est grand. De plus, on peut supposer que pour un nombre d'observations n fixé, le niveau de $\mathcal{R}_\alpha$ sera d'autant plus proche de $1-\alpha$ que l'approximation de F par une fonction linéaire en θ au voisinnage de Θ sera valide.

* à cause des difficultés numériques nous devons chercher une région de confiance approchée pour ψ, notée $\mathcal{R}_{1\alpha'}$, et là encore nous ne contrôlons pas le niveau de cette région.

* enfin, lorsque nous cherchons les extrêmum de $F^{-1}(\psi)$ pour les valeurs de ψ appartenant au contour de $\mathcal{R}_{1\alpha'}$, l'intervalle de confiance ainsi calculé pour x^I est de niveau supérieur à $1-\alpha'$.

Résultats : Calcul d'un intervalle de confiance pour la dose inconnue à l'aide des deux méthodes. On suppose dans tous les cas que le nombre de répétitions de la réponse est égal à 2.

Dosage de la prolactine

Réponse moyenne observée en (cpm)	dose estimée en ng/4 µl	Intervalle de confiance à 95 %	
		1ère méthode	2ème méthode
4200	0,021	[0,011 ; 0,040]	dose nulle
4000	0,043	[0,030 ; 0,060]	[0,021 ; 0,066]
2500	0,32	[0,29 ; 0,36]	[0,28 ; 0,37]
2000	0,51	[0,45 ; 0,58]	[0,45 ; 0,58]
500	3,4	[2,6 ; 4,6]	[2,6 ; 4,8]

La seconde méthode nous indique que la dose minimum significativement non nulle vaut 0,025 ng/5 µℓ et que la dose maximum significativement différente de l'infini est supérieure à 5 ng/5 µℓ. Autrement dit, si le pourcentage de la réponse observée par rapport à une dose nulle d'hormone est supérieur à 94 %, la dose est nulle.

Dosage de la testostérone

Réponse moyenne observée (en cpm)	dose estimée en pg/mℓ	Intervalle de confiance à 95 %	
		1ère méthode	2ème méthode
1200	61	[46 ; 78]	dose nulle
1150	70	[58 ; 83]	[39 ; 88]
800	123	[111 ; 134]	[107 ; 139]
600	165	[151 ; 180]	[147 ; 185]
400	245	[223 ; 270]	[219 ; 277]
70	2180	[1470 ; 3270]	dose infinie

La dose minimum significativement non nulle vaut 70 pg/mℓ ; la dose maximum significativement différente de l'infini vaut 2000 pg/mℓ ; 700 pg/mℓ correspond à un pourcentage de réponse observée de 90 %.

Dosage du Cortisol

Réponse moyenne observée (en cpm)	Dose estimée en $ng/10^{-1}ml$	Intervalle de confiance à 95 %	
		1ère méthode	2ème méthode
2600	0,022	[0,014 ; 0,036]	dose nulle
2500	0,032	[0,024 ; 0,044]	[0,015 ; 0,049]
2000	0,085	[0,076 ; 0,096]	[0,072 ; 0,100]
1400	0,19	[0,17 ; 0,20]	[0,17 ; 0,21]
1200	0,25	[0,23 ; 0,27]	[0,23 ; 0,27]
300	2,55	[2,27 ; 2,88]	[2,27 ; 2,91]

La dose minimum significativement non nulle vaut 0,026 $ng/10^{-1}$ ml (soit un pourcentage de 93 %), la dose maximum significativement différente de l'infini est supérieure à 4 $ng/10^{-1}ml$.

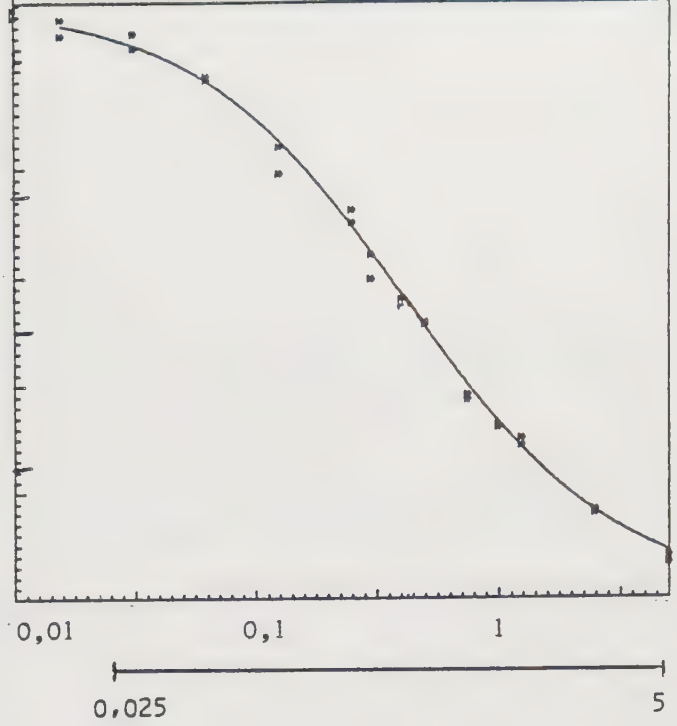
Figure 18

Préparations standards pour les trois dosages étudiés.

* : comptages observés

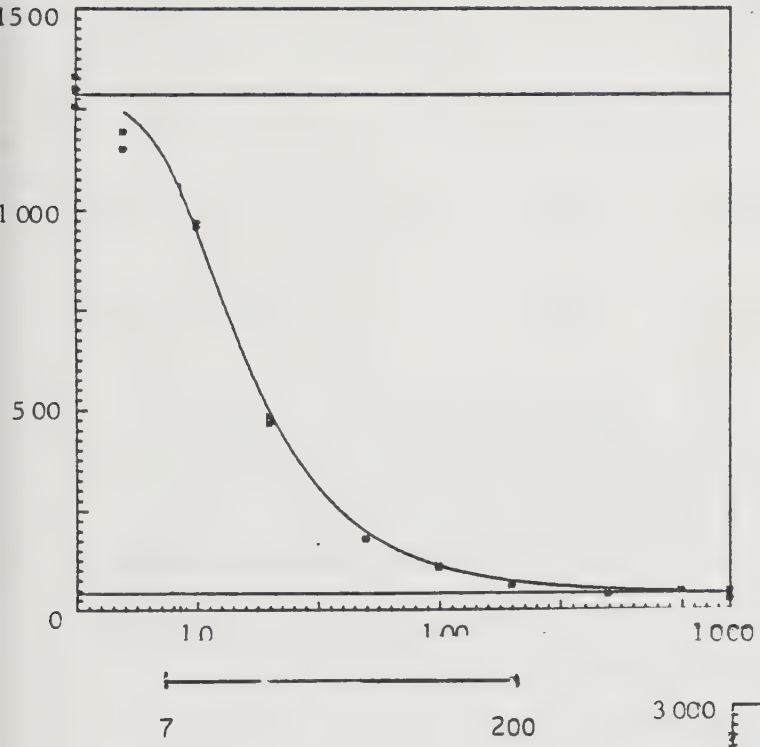
—— : courbe estimée

Le trait horizontal sous le graphe représente l'intervalle des doses où la dose est à la fois significativement non nulle et significativement différente de l'infini selon le calcul de l'intervalle de confiance par la seconde méthode.



↑

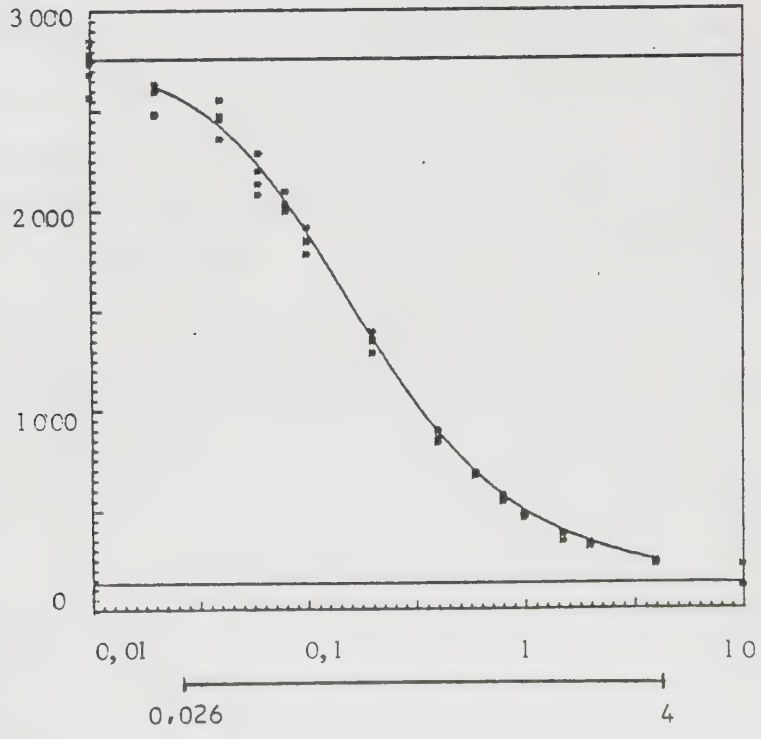
Dosage de la prolactine



← Dosage de la testostérone

Dosage du cortisol

↓



II.6.4. Comment se servir des dilutions ?

Supposons cette fois, que la réponse pour la préparation inconnue étudiée soit mesurée en plusieurs dilutions ($\ell>1$). Nous distinguerons deux cas :

1er cas : le nombre de dilutions est faible, ℓ vaut 2 ou 3. Pour estimer la dose inconnue, nous pouvons utiliser ces dilutions de différentes façons :

- pour chacune des dilutions d_i, i variant de 1 à ℓ, l'estimateur du logarithme de la dose inconnue I, vaut

$$X_i^I = (Y_i^I - A)/B - \text{Log}\ (d_i)$$

où

$$Y_i^I = -\ell n\ ((D-U_{i.}^I)/(U_{i.}^I - C))$$

avec

$$U_{i.}^I = \sum_{j=1,r_i} U_{ij}^I / r_i$$

(pour les notations, consulter le schéma du paragraphe II.2.), les paramètres α,β,d et c ayant été estimés par A,B,D et C à l'aide des réponses observées pour la préparation standard. Nous pouvons alors calculer à l'aide de l'une des deux méthodes précédemment exposées, un intervalle de confiance pour z_i^I, la quantité inconnue d'hormone contenue dans une dilution d_i de préparation I. Bien évidemment, les z_i^I, pour i variant de 1 à ℓ, sont égales et donc les X_i^I sont ℓ estimateurs de la même quantité x^I. Nous pourrions en effectuant une moyenne pondérée des X_i^I calculer un estimateur X^I du logarithme de la dose inconnue d'hormone contenue dans la préparation I. Mais immédiatement, plusieurs questions se posent : Quelle pondération choisir ? Quelles seront les propriétés statistiques de X^I. Comment calculerions nous un intervalle de confiance pour le paramètre x^I ? La méthode paraît peu rigoureuse, et il nous semble plus raisonnable de choisir parmi ces deux ou trois valeurs calculées pour la dose inconnue, celle pour laquelle la réponse observée $U_{i.}^I$ est la plus proche du milieu des deux asymptotes (C+D)/2. A cet endroit de la courbe standard, la pente est plus forte, et la précision de l'estimation est meilleure.

Ainsi, nous venons de proposer une première méthode d'estimation de la dose inconnue.

- une seconde méthode consiste à prendre en compte l'ensemble des observations effectuées pour la préparation inconnue I. En effet, si la condition de similitude est vérifiée, et si $F^I(z)$ est la fonction de règression pour la préparation inconnue I (voir paragraphe I.3.), alors nous pouvons écrire l'équation suivante pour toute valeur de z : $F^I(z) = F(z^I z)$, où z^I est la dose inconnue d'hormone contenue dans la préparation inconnue I. Nous en déduisons , $F^I(z) = c + (d-c)/(1+\exp-(\alpha^I+\beta \text{ Logz}))$ où $\alpha^I = \alpha + \beta \text{ Log } z^I$. Ainsi, le logarithme de la dose inconnue, x^I, vaut $(\alpha^I-\alpha)/\beta$. Pour les raisons que nous exposons dans l'annexe B, 2ème cas, les paramètres α,β,c,d sont estimés par A,B,C,D à l'aide des seules observations effectuées pour la préparation standard. Puis nous supposons que $F^I(z) = C + (D-C)/(1+\exp-(\alpha^I+B \text{ Log}(z)$ où C,D et B sont considérés comme des constantes. Nous en déduisons une estimation A^I de α^I à l'aide des seules observations effectuées pour la préparation inconnue I. Ainsi, x^I est estimé par $X^I = (A^I-A)/B$. Comment allons nous calculer un intervalle de confiance pour x^I ? Nous pouvons soit appliquer le théorème de Fieller au rapport $(\alpha^I-\alpha)/\beta$ (paragraphe II.6.2.), soit appliquer la seconde méthode (paragraphe II.6.3.) au vecteur $\psi = (\alpha^I, c,d,\alpha,\beta)$ et à la fonction : $\psi \longrightarrow (\alpha^I-\alpha)/\beta$. En fait, le problème réside dans l'estimation des corrélations de A^I avec A,B,C,D. La méthode d'estimation de A^I revient en effet à supposer A^I indépendant de C,D et B ce qui est bien évidemment faux. Aussi, nous pensons plus raisonnable et plus rigoureux d'utiliser la méthode précédemment exposée pour estimer la dose inconnue et calculer un intervalle de confiance quand nous disposons de deux ou trois dilutions.

2ème cas : les dilutions sont suffisamment nombreuses et bien réparties sur l'intervalle d'étude pour estimer les paramètres de la courbe de réponse pour la préparation inconnue.

Supposons que pour la préparation inconnue I, ℓ soit supérieur à 5. Supposons la condition de similitude vérifiée, alors $F^I(z)$ la fonction de régression pour la préparation inconnue I, s'écrit :

$$F^I(z) = c + (d-c)/(1+\exp-(\alpha^I+\beta \text{ Log } z)) \quad \text{où}$$

$\alpha^I = \alpha + \beta \text{ Log } z^I$. Les paramètres α,β,c,d et α^I sont estimés par A,B,C,D et A^I

à l'aide de toutes les observations effectuées pour la préparation standard et pour la préparation inconnue I (voir 1er cas de l'annexe B). Nous sommes revenus au plan d'expérience classique d'un dosage biologiques (voir chapitre I). Le logarithme de la dose x^I, est comme précédemment estimé par $X^I = (A^I-A)/B$, et nous sommes en mesure d'estimer la matrice de variance-covariance des estimateurs A,B,C,D et A^I. L'application de la méthode proposée au paragraphe II.6.3., au vecteur $\psi = (\alpha^I,c,d,\alpha,\beta)$ et à la fonction $\psi \rightarrow (\alpha^I-\alpha)/\beta$ permet de calculer un intervalle de confiance pour x^I.

II.7. PLAN D'EXPERIENCE .-

Le but de l'analyse statistique du dosage radioimmunologique est d'obtenir la meilleure estimation possible de l'intervalle de confiance autour des doses inconnues. En même temps nous souhaitons que pour une dose donnée, la longueur de l'intervalle de confiance soit la plus faible possible. A priori, ces deux problèmes n'admettent pas les mêmes solutions. Par exemple, nous avons vu au paragraphe II.6. que passer de 2 à 4 répétitions de la réponse pour une préparation inconnue diminue considérablement la longueur de l'intervalle de confiance estimé, toutes choses restant égales par ailleurs. Malgré tout, si pour calculer les bornes de l'intervalle de confiance nous appliquons la formule du théorème de Fieller, nous voyons que les résultats vont non seulement dépendre du choix de la courbe standard (symétrique ou non) et des estimations de ses paramètres, mais aussi des variances et covariances estimées des estimateurs.

Préoccupons nous, dans un premier temps de la recherche du plan d'expérience permettant d'estimer au mieux les paramètres de la courbe standard. Un tel plan d'expérience est dit optimal. Le problème des plans d'expérience optimaux a été étudié dans le cas où la variable aléatoire observée a une espérance dépendant linéairement d'un ou plusieurs paramètres et une variance constante (FEDOROV, 1972). Par exemple dans le cas simple de la régression linéaire $Y = \theta_1 + \theta_2 x + E$ où l'erreur E est supposée d'espérance nulle et de variance constante, on peut déterminer analytiquement un plan d'expérience optimal. Nous ne ferons pas la démonstration ici et nous nous contenterons d'exposer les résultats dans le paragraphe II.7.1. Mais pour notre modèle logistique non linéaire que pourrons nous en déduire ? (paragraphe II.7.2.). Toutefois, ne perdons pas de vue

le but de notre analyse statistique et assurons nous qu'un plan d'expérience optimal pour les paramètres de la courbe standard l'est aussi pour la dose inconnue à estimer (paragraphe II.7.3.).

Cette question a déjà fait l'objet de plusieurs articles : CHANG et al. (1975) se situent dans le modèle linéaire en estimant la courbe standard à l'aide de la méthode logit-log. Ils établissent que la dose inconnue est estimée avec une grande précision si les règles suivantes sont respectées :

- les réponses observées pour les différentes doses de la courbe standard sont bien réparties sur un intervalle variant de 20 % à 80 % de la réponse estimée pour une dose nulle d'hormone.

- le nombre de doses permettant l'estimation de la courbe standard vaut au moins huit.

- le nombre de répétitions de la réponse en chaque dose est deux ou plus.

- les réponses observées pour les préparations inconnues sont dans l'intervalle variant de 20 % à 80 % de la réponse estimée pour une dose nulle d'hormone.

De plus, ils estiment les paramètres de la courbe standard (notés α et β dans le paragraphe II.3.5.) à l'aide d'une régression linéaire simple et d'une régression linéaire pondérée. Sous les conditions énumérées ci-dessus, ils assurent que les deux méthodes sont équivalentes. Leur analyse ne fait appel à aucun développement théorique et ne repose que sur des comparaisons numériques. HEALY (1972) s'appuie sur les travaux de BOX (1968) et énonce que le meilleur plan d'expérience pour estimer la courbe standard consiste à mesurer la réponse en seulement quatre doses distinctes (y compris les doses nulle et infinie).

Nous allons tenter d'énoncer quelques principes qui permettront au biologiste de déterminer au mieux un plan d'expérience.

II.7.1. Cas particulier : la régression linéaire

Supposons que la réponse Y soit une fonction linéaire de la dose x. $Y_{ij} = \theta_1 + \theta_2 x_i + E_{ij}$. Les E_{ij} sont supposés indépendants, identiquement distribués, d'espérance nulle et de variance constante égale à σ^2 ; j varie de 1 à n_i pour i variant de 1 à k. (k est au moins égal à 2). Nous cherchons le plan d'expérience $\mathcal{P}$, c'est-à-dire l'ensemble des points $(n_1,x_1,n_2,x_2,\ldots,n_k,x_k)$,

tel que la relation entre la dose et la réponse soit estimée le mieux possible sachant qu'elle est linéaire.

Nous pouvons choisir différents critères. Par exemple nous pouvons chercher à minimiser une caractéristique de la matrice de variance-covariance des estimateurs de θ_1 et θ_2 comme son déterminant. Soit θ le vecteur des paramètres, Y le vecteur des observations, E le vecteur de l'erreur résiduelle.

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \qquad Y' = (Y_{11} \cdots Y_{1n_1} \cdots Y_{k_1} \cdots Y_{kn_k})$$

$$E' = (E_{11} \cdots E_{1n_1} \cdots E_{k1} \cdots E_{kn_k}), \text{ (' signifiant transposé de)}.$$

Ainsi nous pouvons écrire le modèle de la façon suivante

$Y = \underset{\sim}{X}(\mathcal{P})\ \theta + E$ où $\underset{\sim}{X}(\mathcal{P})$ appelée matrice du plan d'expérience est définie par

$\underset{\sim}{X}(\mathcal{P}) = \underset{\sim}{A}\ \underset{\sim}{\mathcal{X}}$ avec $\underset{\sim}{\mathcal{X}} = \begin{pmatrix} \mathbb{1}_k & x_1 \\ & x_2 \\ & \vdots \\ & x_k \end{pmatrix}$ où $\mathbb{1}_k$ désigne le vecteur composé de k 1 et $\underset{\sim}{A}$ est la matrice symétrique égale à

$$\underset{\sim}{A} = \begin{pmatrix} \mathbb{1}_{n_1} & & & \\ 0_{n_2} & \mathbb{1}_{n_2} & & \\ \vdots & \ddots & \ddots & \\ 0_{n_k} & 0_{n_{k-1}} & & \mathbb{1}_{n_k} \end{pmatrix}$$

où 0_{n_i} est le vecteur composé de n_i 0.

La théorie du modèle linéaire nous permet d'énoncer que la variance de Θ vaut $\sigma^2\ (\underset{\sim}{X}'(\mathcal{P})\ \underset{\sim}{X}(\mathcal{P}))^{-1}$. Ainsi nous chercherons le plan $\mathcal{P}$ qui rendra maximum le déterminant de $X'(\mathcal{P})\ X(\mathcal{P})$. Si nous supposons que x varie entre -1 et +1 (si ce n'est pas le cas, il suffira de transformer x pour s'y ramener), nous pouvons facilement montrer que le plan d'expérience $\mathcal{P}$ qui minimise la variance-covariance

des estimateurs est $\mathcal{P} = (\frac{n}{2}, -1, \frac{n}{2}, +1)$ c'est-à-dire le plan d'expérience formé de n/2 répétitions de la réponse pour la dose x=-1 et n/2 répétitions pour la dose x=+1.

Supposons maintenant les E_{ij} non identiquement distribués :
Var $E_{ij} = \sigma^2 \gamma_i$ où les γ_i qui dépendent de x_i sont supposés certains (non aléatoires) et connus a priori. Ainsi nous écrivons Var $E = \sigma^2 \underset{\sim}{\Gamma}$ où $\underset{\sim}{\Gamma}$ est la matrice définie par

$$\underset{\sim}{\Gamma} = \begin{pmatrix} \gamma_1 \underset{\sim}{\mathbb{I}}_{n_1} & & & \\ & \gamma_2 \underset{\sim}{\mathbb{I}}_{n_2} & & \\ & & \ddots & \\ & & & \gamma_k \underset{\sim}{\mathbb{I}}_{n_k} \end{pmatrix}$$

où $\underset{\sim}{\mathbb{I}}_{n_1}$ désigne la matrice identité à n_1 lignes et n_1 colonnes.

La matrice de variance-covariance de Θ vaut $\sigma^2 (\underset{\sim}{X}'(\mathcal{P})\, \Gamma^{-1}\, \underset{\sim}{X}(\mathcal{P}))^{-1}$.
Comme précédemment cherchons $\mathcal{P}$ qui rendra maximum le déterminant de $\underset{\sim}{X}'(\mathcal{P})\, \Gamma^{-1}\, \underset{\sim}{X}(\mathcal{P})$.

Il n'existe pas comme dans le cas précédent de solution analytique simple. Aussi, on utilise un algorithme numérique (METZER et al., 1976) dont la solution est souvent très proche du plan $\mathcal{P}$: $\left\{ \frac{n\,\gamma_1}{\gamma_1+\gamma_2}, -1, \frac{n\,\gamma_2}{\gamma_1+\gamma_2}, +1 \right\}$.

Ainsi, sachant que la relation entre la dose et la réponse est linéaire sur l'intervalle $[-1,+1]$, $Y = \theta_1 + \theta_2 x$, la plan d'expérience qui permettra d'estimer au mieux les paramètres θ_1 et θ_2 consiste à mesurer la réponse en deux doses distinctes le plus éloignées possible, $x_1=-1$ et $x_2=+1$. Le nombre de répétitions est proportionnel à la variance de la réponse en ces doses.

II.7.2. Le modèle logistique à 4 paramètres

Reprenons les notations du paragraphe II.6.3.

Supposons que nous estimions la matrice de variance-covariance de θ, par

$$V_\theta = -\left(\frac{\partial^2 \ell n L_2}{(\partial\theta)^2} (\theta) \right)^{-1} \qquad \text{(BARD, 1974)}$$

Suite au paragraphe précédent, notre première idée est de chercher à minimiser le déterminant de V_θ en fonction de $\mathcal{P} = \{n_1, x_1, \ldots, n_k, x_k\}$. Mais le problème est en fait bien différent, d'une part parce que le déterminant de V_θ est une fonction très compliquée de $\mathcal{P}$ (s'il faut s'en convaincre , voir dans l'annexe B comment s'écrit la vraisemblance) et qu'il est loin d'être évident de programmer un algorithme numérique qui fasse un tel calcul, d'autre part parce que V_θ dépend de θ. Ainsi, le plan d'expérience optimal s'il existe, dépend de la vraie valeur des paramètres. Toutefois, considérons plusieurs dosages, de la même hormone effectués dans le même laboratoire à l'aide de la même méthode. Généralement, la modélisation que nous ferons de la réponse pour la préparation standard ne sera pas fondamentalement différente d'un dosage à l'autre. Par exemple pour les dosages de la testostérone dont nous avons étudié un exemple, la variabilité de la variance des réponses d'un dosage à l'autre ne varie pas, de plus la courbe reste dissymétrique (voir les résultats du tableau 3 où une valeur de h égale à 1,2 est toujours convenable, γ est souvent proche de 0,5). Aussi, les quelques principes que nous énoncerons pour un plan d'expérience $\mathcal{P}$ d'un dosage particulier, nous permettront d'améliorer la qualité de l'estimation des paramètres de la courbe standard du dosage suivant.

Tableau 3

Estimation des paramètres de la courbe standard pour différents dosages de la testostérone. Dans tous les cas h vaut 1,2.

Date du dosage	γ	Valeurs calculées pour			
		C	D	A	B
26/11/80	0,4	38	1609	6,12	-6,44
2/12/80	0,5	26	924	3,58	-4,34
13/10/81	0,5	29	1006	4,58	-5,16
28/10/81	1	33	1274	3,30	-3,00

Intéressons nous à quelques paramètres dont il est simple de comprendre la signification.

* les coefficients de variation

Le coefficient de variation du paramètre c, par exemple, est le rapport de l'écart-type estimé pour c sur l'estimation de c. La comparaison des coefficients de variation calculés des différents paramètres, nous indique si les données contribuent de la même façon à l'estimation de chacun des paramètres.

1er exemple : dosage de la prolactine

Reprenons les résultats que nous avons obtenus au paragraphe II.5. Dans un premier temps, nous avons estimé la courbe standard en estimant les 4 paramètres c,d,α et β.

	Coefficient de variation calculé
c	60
d	0,011
α	0,101
β	0,052

Il apparaît avec évidence que le coefficient de variation pour c est très grand devant les coefficients de variation pour d,α,β. Ainsi, les observations n' apportent aucune information sur c. Pour ce dosage nous avions supposé a priori que c était nul, il était donc raisonnable de l'éliminer de l'ensemble des paramètres. D'où les résultats suivants :

	Coefficient de variation calculé
d	0,009
α	0,033
β	0,028

2ème exemple : dosage de la testostérone

Rappelons que h vaut 1,2, que la courbe est dissymétrique avec γ=0,3.

	Coefficient de variation calculé
c	0,055
d	0,016
α	0,049
β	0,033

d semble favorisé par rapport aux autres.

3ème exemple : dosage du cortisol

h vaut 1,7, la courbe est dissymétrique avec γ=0,6

	Coefficient de variation calculé
c	0,020
d	0,008
α	0,007
β	0,011

Le paramètre c semble défavorisé par rapport aux autres.

A l'aide du coefficient de variation, nous ne recueillons qu'une information globale. Nous ne pouvons pas par exemple répondre à la question suivante : où doit-on ajouter des observations pour obtenir une meilleure estimation des paramètres de la courbe standard ?

* les coefficients de sensibilité

Intéressons nous par exemple à l'estimation du paramètre β et étudions les variations de $F(\theta,x)$ par rapport aux variations de β seul, en fonction de x. Nous allons faire notre étude au voisinage de θ. La variation de $F(\theta,x)$ par rapport aux variations de β vaut $\frac{\partial F(\theta,x)}{\partial \beta}$. Les valeurs de x pour lesquelles $\frac{\partial F(\theta,x)}{\partial \beta}$ est grand en valeur absolue sont appelées les entrées les plus sensibilisantes pour β. Des observations nouvelles en ces doses permettront une estimation plus précise de β. Les figures qui suivent montrent les courbes de la valeur absolue des dérivées partielles de $F(\theta,x)$ par rapport à c,d,α et β en fonction de x. Pour les trois exemples déjà traités, ces dérivées sont calculées pour la valeur estimée de θ. De plus pour tenir compte de la variabilité de la variance nous avons pondéré chaque point par $\phi_h\left[F(\theta,x)\right]^{1/2}$. Ainsi si $\hat{\theta}$ est la valeur calculée de θ les figures qui suivent représentent les courbes de $\frac{1}{\phi_h\left[F(\hat{\theta},x)\right]} \frac{\partial F(\hat{\theta},x)}{\partial c}$ en fonction de x, et de même par rapport à d,α et β.

Pour le dosage de la prolactine (figure 19), les entrées les plus sensibilisantes pour l'asymptote supérieure sont évidemment les faibles doses, plus précisément les doses comprises entre 0 et 0,125 ng/5 µℓ. La dérivée de F par rapport à α présente un maximum pour x environ égal à 0,4 ng/5 µℓ. La valeur absolue de la dérivée de F par rapport à β présente deux maximum pour des doses approximativement égales à 0,125 et 5 ng/5 µℓ.

Pour les dosages de la testostérone (figure 20) et du cortisol (figure 21), les allures des courbes restent inchangées malgré une dissymétrie de la fonction de réponse F et la variabilité de la variance. Pour une meilleure estimation de l'asymptote inférieure c, il semblerait opportun en ce qui concerne le dosage du cortisol d'ajouter de fortes doses, de l'ordre de 20 ng/100 µℓ.

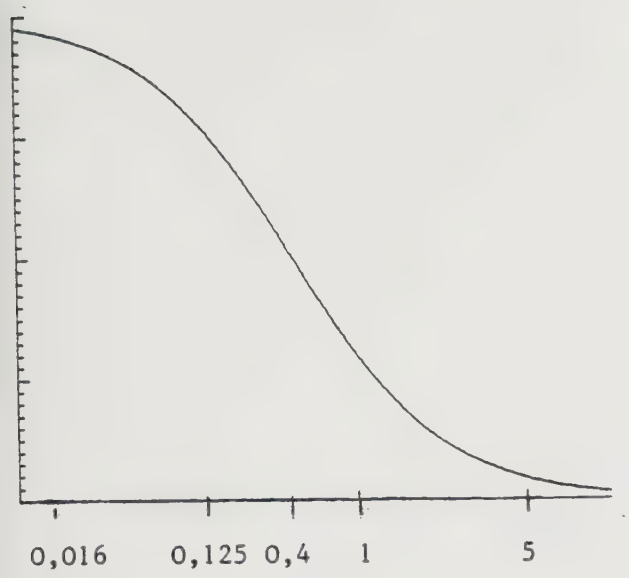


Figure 19.a.

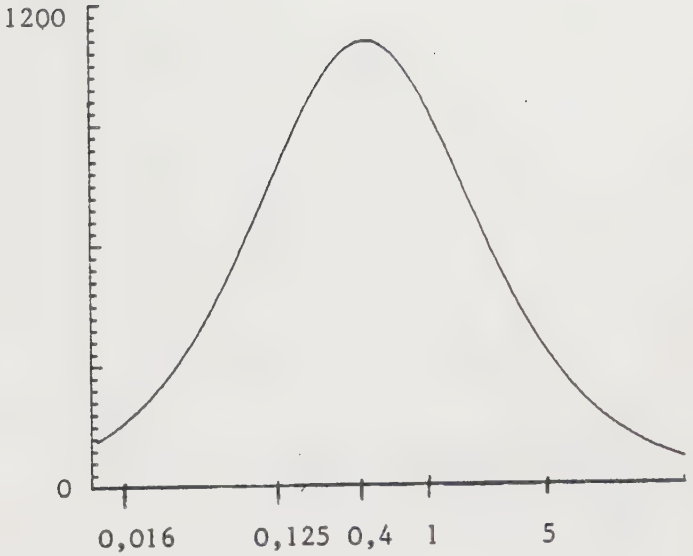


Figure 19.b.

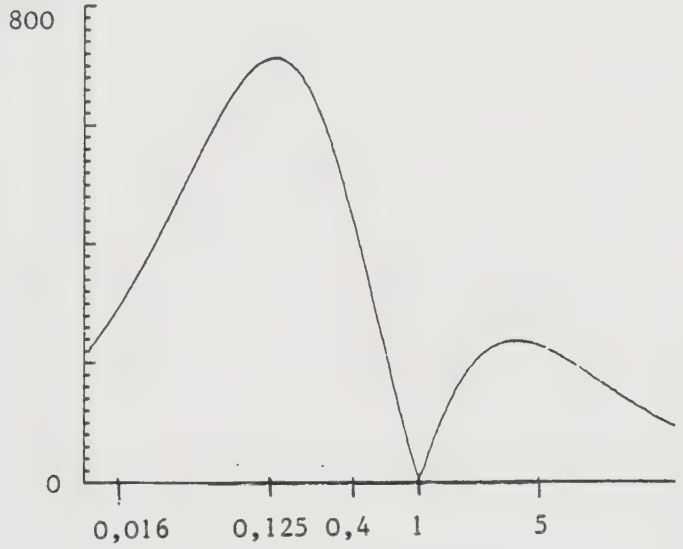


Figure 19.c.

Figure 19 : Dosage de la prolactine

Figure 19.a. : Graphe de $\frac{\partial}{\partial d} F(\hat{\theta},x)$ en fonction de x.

Figure 19.b. : Graphe de $\frac{\partial}{\partial \alpha} F(\hat{\theta},x)$ en fonction de x

Figure 19.c. : Graphe de $\frac{\partial}{\partial \beta} F(\hat{\theta},x)$ en fonction de x

Figure 20 : Dosage de la testostérone

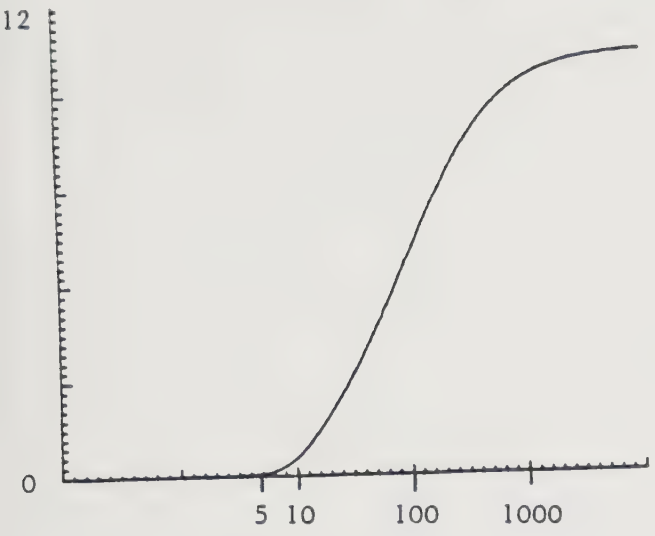


Figure 20.a.

Graphe de $\frac{1}{\sqrt{F(\hat{\theta},x)^{1,2}}} \frac{\partial}{\partial c}(F(\hat{\theta},x)$

en fonction de x.

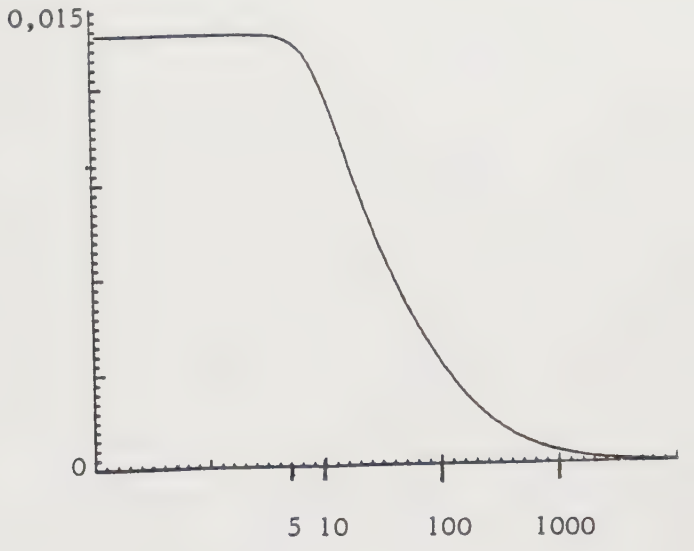


Figure 20.b.

Graphe de $\frac{1}{\sqrt{F(\hat{\theta},x)^{1,2}}} \frac{\partial}{\partial d} F(\hat{\theta},x)$

en fonction de x.

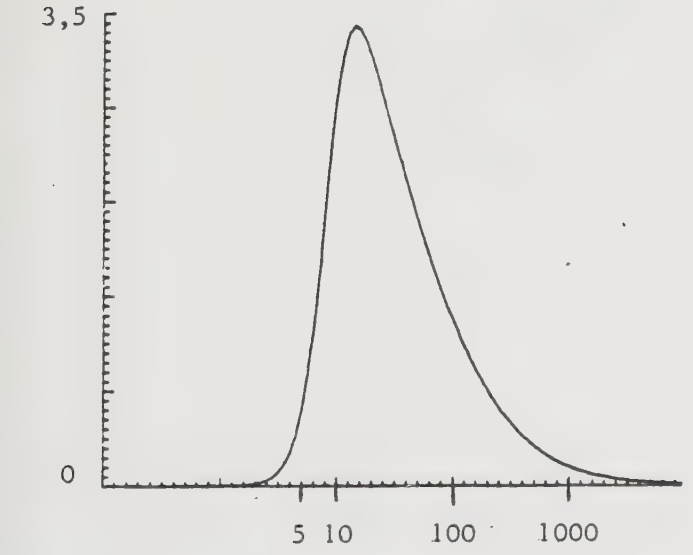


Figure 20.c.

Graphe de $\frac{1}{\sqrt{F(\hat{\theta},x)^{1,2}}} \frac{\partial}{\partial \alpha} F(\hat{\theta},x)$

en fonction de x.

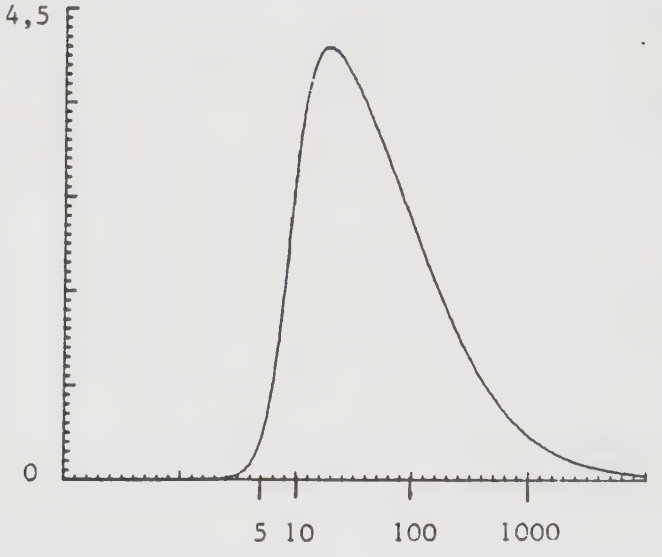


Figure 20.d.

Graphe de $\frac{1}{\sqrt{F(\hat{\theta},x)^{1;2}}} \frac{\partial}{\partial \beta} F(\hat{\theta},x)$

en fonction de x.

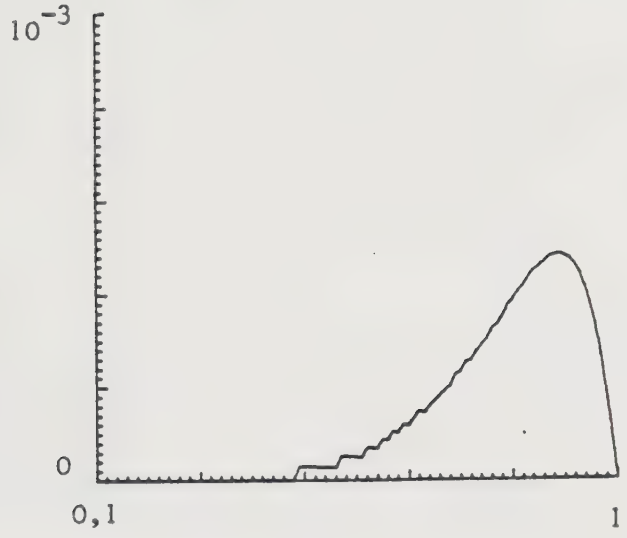


Figure 20.e.

Graphe de $\frac{1}{\sqrt{F(\hat{\theta},x)^{1,2}}} \frac{\partial}{\partial\beta} F(\hat{\theta},x)$ en fonction du logarithme de la dose pour des doses comprises entre 0,1 et 1 ng pour 0,1 ml.

Figure 21 : Dosage du cortisol

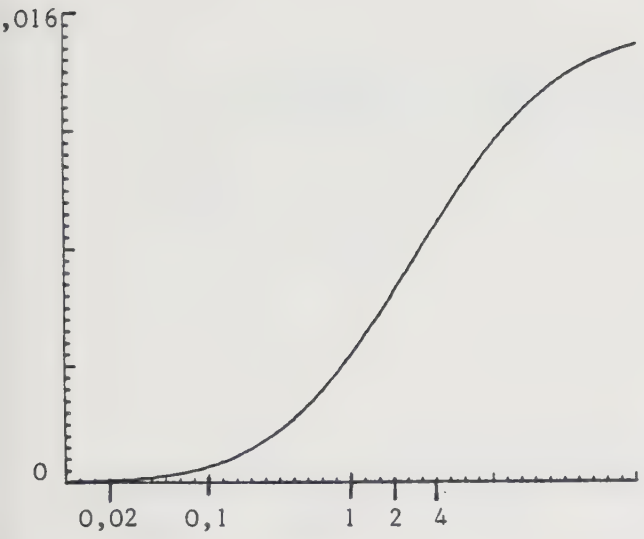


Figure 21.a.

Graphe de $\frac{1}{\sqrt{F(\hat{\theta},x)^{1,7}}} \frac{\partial}{\partial c} F(\hat{\theta},x)$ en fonction de x.

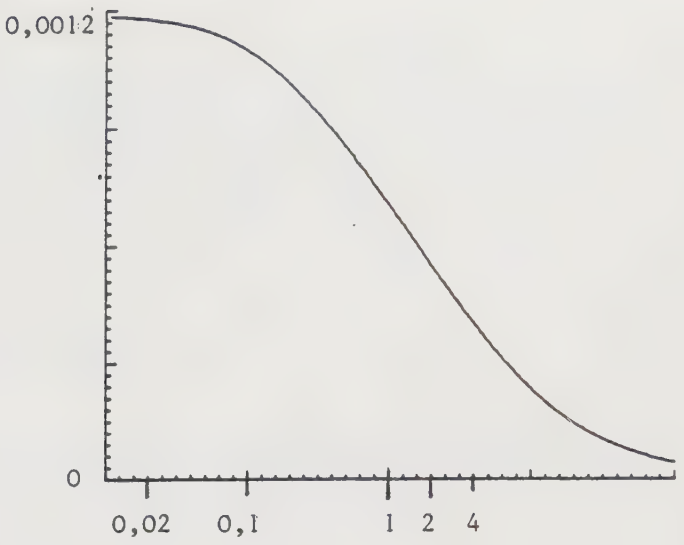


Figure 21.b.

Graphe de $\frac{1}{\sqrt{F(\hat{\theta},x)^{1,7}}} \frac{\partial}{\partial d} F(\hat{\theta},x)$ en fonction de x.

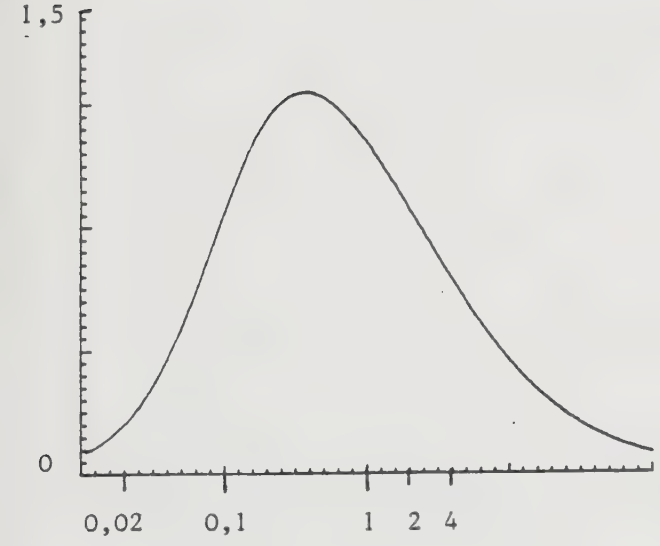


Figure 21.c.

Graphe de $\frac{1}{\sqrt{F(\hat{\theta},x)^{1,7}}} F(\hat{\theta},x)$ en fonction de x.

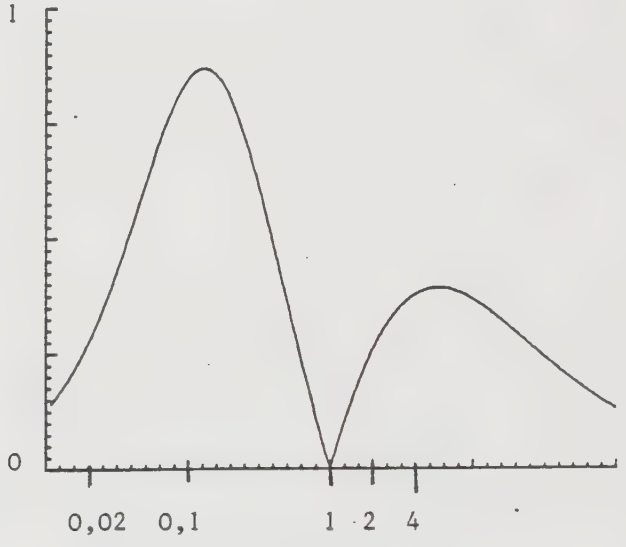


Figure 21.d.

Graphe de $\frac{1}{\sqrt{F(\hat{\theta},x)^{1,7}}} \frac{\partial}{\partial \beta} F(\hat{\theta},x)$ en fonction de x.

* Essayons de trouver un critère qui prenne en compte globalement toute l'information obtenue pour chacun des paramètres. Supposons que $\hat{\theta}$ soit la valeur calculée pour θ. La linéarisation de $F(\theta,x)$ dans un voisinnage de $\hat{\theta}$ conduit à l'approximation

$$V_{\hat{\theta}} = \frac{s^2}{n-p} \; (\underset{\sim}{X}' \; \underset{\sim}{X})^{-1}$$

où $\underset{\sim}{X}$ est la matrice du plan d'expérience formée de n lignes et p colonnes, p étant le nombre de paramètres. La ième ligne de $\underset{\sim}{X}$ vaut

$$\left(\frac{\partial F(\hat{\theta},x_i)}{\partial c} \quad \frac{\partial F(\hat{\theta},x_i)}{\partial d} \quad \frac{\partial F(\hat{\theta},x_i)}{\partial \alpha} \quad \frac{\partial F(\hat{\theta},x_i)}{\partial \beta} \right.$$

s^2 est la valeur calculée de $\sum_{i,j} (U_{ij} - F(\theta,x_i))^2$, la somme des carrés résiduelle. Nous nous rapprochons ainsi de la solution du problème que nous nous sommes posés au paragraphe II.7.1. et nous pouvons espérer que la minimisation du déterminant de $V_{\hat{\theta}}$ soit numériquement réalisable. Un programme écrit par METZLER et al. (1976) permet de déterminer un plan $\mathcal{P}$ qui maximise le déterminant de la matrice $\underset{\sim}{X}' \; \underset{\sim}{X}$.

Nous avons fixé n égal à 12 et nous avons cherché le plan d'expérience qui maximise le déterminant de la matrice $\underset{\sim}{X}' \; \underset{\sim}{X}$ en supposant c=0, d=3000, α=0 et β=-1. Nous avons trouvé que le plan d'expérience optimal était composé de la façon suivante $(3,-\infty;3,-1;3,+1;3,+\infty)$. Ce qui signifie 3 répétitions de la réponse pour chacune des doses suivantes: les doses nulles et infinie, les doses de logarithme égal à +1 et -1. Si n vaut 16, nous retrouvons les mêmes doses de base et le plan vaut $(4,-\infty;4,-1;4,+1;4,+\infty)$. Si nous posons β = -0,5, le plan $\mathcal{P}$ vaut $(3,-\infty;3,-2;3,+2;3,+\infty)$. Si β vaut -5, le plan $\mathcal{P}$ vaut $(3,-\infty;3,-0,2;3,-0,2;3,+\infty)$. Si β vaut -1 et α vaut -1 (alors, l'abscisse du

point d'inflexion n'est plus égale à 0 mais à -1), le plan $\mathcal{P}$ vaut $(3,-\infty;3,-2;3,0;3,+\infty)$. Ainsi si n est un multiple de 4, le plan d'expérience optimal est constitué de 4 doses pour lesquelles la réponse doit être répétée n/4 fois. Il s'agit de doses nulle et infinie, et des doses dont les logarithmes se trouvent de part et d'autre du point d'inflexion à une distance inversement proportionnelle à β.

Ce résultat n'est valable que dans le cas où la variance des réponses est constante. Si cela n'est pas le cas, nous pouvons considérer le déterminant de la matrice $\underset{\sim}{X}' \underset{\sim}{X}$ où cette fois ci la ième ligne de $\underset{\sim}{X}$ s'écrit

$$\frac{1}{\sqrt{\phi_h[F(\hat{\theta},x_i)]}} \left(\frac{\partial F(\hat{\theta},x_i)}{\partial c} \quad \frac{\partial F(\hat{\theta},x_i)}{\partial d} \quad \frac{\partial F(\hat{\theta},x_i)}{\partial \alpha} \quad \frac{\partial F(\hat{\theta},x_i)}{\partial \beta} \right)$$

Nous n'avons pas programmé un tel calcul, nous pouvons toutefois espérer trouver un résultat similaire à celui que nous avons énoncé au paragraphe précédent où les doses de base restent identiques à celles déterminées dans le cas où la variance est constante et où le nombre de répétitions est proportionnel à la variance des réponses en ces doses. Une étude reste également à faire dans le cas où la fonction de réponse n'est pas symétrique.

Toutefois, si nous choisissons un plan d'expérience avec quatre doses de base, implicitement nous ne remettons pas en cause la modélisation de l'espérance de la réponse par la logistique. Ce n'est en pratique pas raisonnable. Pour chaque courbe standard étudiée, nous devons être en mesure de détecter un écart au modèle ; et ceci sera d'autant mieux réalisé que le nombre de doses de base sera grand, et que ces doses seront bien réparties sur l'intervalle d'étude.

II.7.3. Conclusion

Le choix du plan d'expérience sera différent, selon s'il s'agit de mettre au point une nouvelle technique de dosage, ou d'un dosage de routine. Dans le premier cas, nous ne connaissons pas la forme de la courbe : n'hésitons pas à choisir un nombre important de doses bien réparties sur tout l'intervalle d'étude. Dans le second cas, si d'un dosage à l'autre la courbe présente la même

allure, à l'aide des coefficients de sensibilité, déterminons pour chacun des paramètres, les entrées les plus sensibilisantes.

Il n'est pas opportun d'affiner d'avantage l'analyse. D'une part nous sommes freinés par la complexité numérique de l'étude, d'autre part, rien ne nous permet d'affirmer qu'une meilleure estimation de la courbe standard entraîne une amélioration sensible de l'estimation des doses inconnues. Certes, il est indispensable de chercher à estimer avec le maximum de précision la réponse pour la préparation standard ; mais en même temps, il est souvent possible d'agir sur les doses inconnues en effectuant des dilutions ou en augmentant le nombre de répétitions de la réponse.

II.8. PROBLEME DES DOSAGES AVEC "EXTRACTION" .-

II.8.1. Exposé du problème à l'aide d'un exemple : le dosage de la testostérone.

Au paragraphe II.4., nous avions déjà soulevé le problème des différentes méthodes de dosage qui étaient employées pour un même dosage suivant que la quantité d'hormone à doser est connue ou non. Cette pratique étant courante pour de nombreuses hormones (notamment les stéroides), il est utile de réfléchir aux conséquences statistiques.

Supposons par exemple que l'on veuille doser la testostérone contenue dans un plasma. Le problème réside dans le fait que le plasma contient d'autres substances que la testostérone et qu'il est indispensable d'isoler la testostérone avant de la mettre en compétition avec de l'hormone chaude pour son anticorps spécifique. Aussi, on ajoute au plasma un solvant, dont le rôle est de capter la testostérone.

A ce stade, intervient une première source d'erreur : il se peut que le solvant ne capte pas toute la testostérone. Après récupération du solvant contenant la testostérone, puis évaporation, seule la testostérone reste adsorbée aux paroiesdu verre. Celle-ci sera récupérée après ajout d'une solution "tampon" ; une agitation vigoureuse permet alors à la testostérone de se décoller des paroies du verre et de se diluer dans le solvant. Une deuxième source d'erreur apparaît alors : la testostérone reste accrochée au verre et ne se dissout pas dans la solution tampon. Encore une fois, on a "perdu" de la testostérone. Afin de savoir si ces manipulations ont été bien conduites, on mesure un paramètre

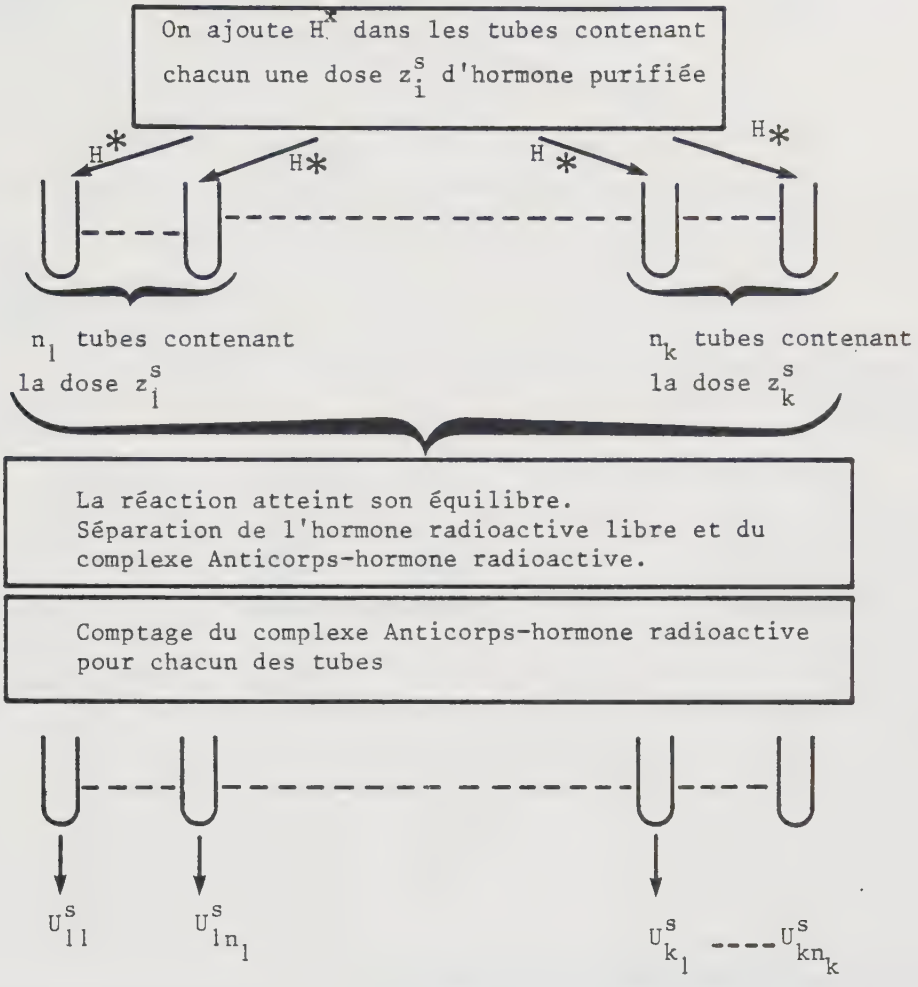
supplémentaire appelé <u>récupération</u> ou encore pourcentage d'extraction. Bien entendu, ces difficultés ne se rencontrent que pour le dosage des préparations inconnues. En effet, les doses de la préparation standard étant constituées de différentes dilutions de l'hormone purifiée, le problème ne se pose pas.

Supposons donc que les k doses permettant l'estimation de la courbe standard soient dosées par une méthode directe ; c'est-à-dire, l'hormone radioactive H^* est ajoutée dans chacun des tubes, puis l'équilibre de la réaction étant atteint, après séparation de l'hormone radioactive liée à l'anticorps et de l'hormone radioactive restée libre, on mesure la quantité d'hormone radioactive liée à l'anticorps. Pour les préparations inconnues, l'hormone radioactive H^* n'est pas directement introduite dans les solutions ; celles-ci subissent auparavant les différentes manipulations que nous venons de décrire. Pour tenter d'évaluer la quantité d'hormone perdue, on mesure la récupération de la façon suivante :
on introduit dans chaque solution une quantité h^* d'hormone radioactive. Après ajout de h^* et manipulations, nous allons partager la solution obtenue en L+1 solutions, où $L = \sum_{i=1}^{\ell} r_i$ est le nombre d'observations de la réponse pour la préparation inconnue I. Dans L d'entre-elles nous allons ajouter comme pour la préparation standard, l'hormone radioactive H^* et mesurer la quantité d'hormone radioactive liée à l'anticorps. En même temps nous allons mesurer la radioactivité contenue dans la dernière solution, qui corrigée par le volume considéré doit être égale à la radioactivité initialement introduite h^*. Si μ^* est le comptage radioactif de h^* dans un volume v, si μ_I^* est le comptage radioactif de h^* dans un volume v_I de la solution obtenue après extraction et évaporation, la récupération p_I de la préparation inconnue I est définie par $p_I = \frac{v}{v_I} \frac{\mu_I^*}{\mu^*}$. Si tout s'est bien passé p_I vaut 1. Si p_I est inférieur à 1, cela signifie qu'une partie de la préparation initiale a été perdue lors de manipulations.

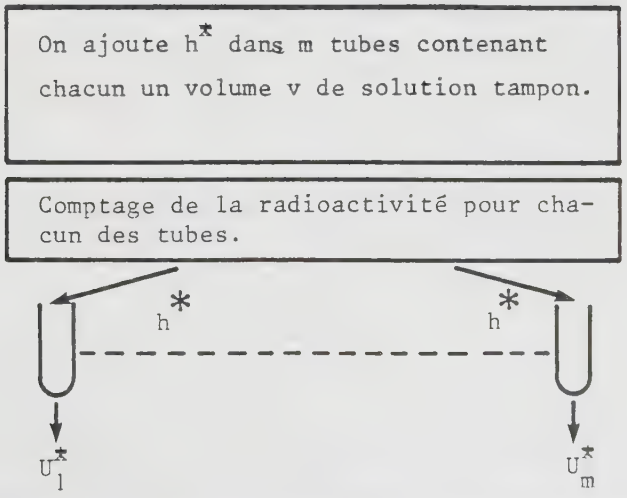
Certains expérimentateurs mesurent également un autre paramètre, appelé le blanc, qui permet de vérifier qu'à la suite des diverses manipulations, la quantité d'hormone mesurée dans une solution "tampon" (solution où la quantité d'hormone est nulle) est nulle.

Schéma d'un dosage à méthode directe pour la préparation standard et à méthode avec "extraction" pour les préparations inconnues.

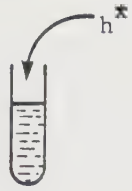
Préparation standard



Estimation de h^*



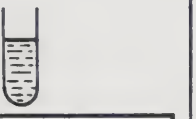
Estimation de la dose d'hormone contenue dans la préparation inconnue I.



On ajoute h^* dans un volume $\mathcal{V}$ de préparation inconnue

extraction et évaporation

On obtient une solution $\mathcal{S}$ dans laquelle on a logiquement récupéré la totalité de h^*.

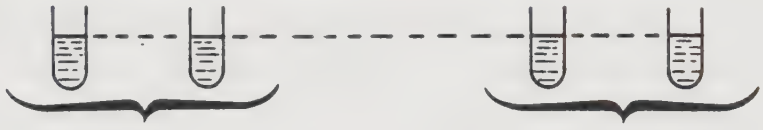


volume v_I de solution $\mathcal{S}$

Comptage de la radioactivité

U^*_I

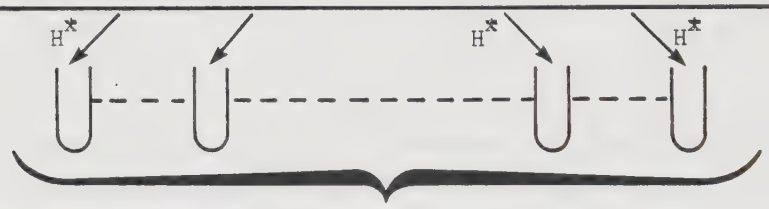
si on veut doser I à l'aide de ℓ dilutions d_i avec pour chacune r_i répétitions de la réponse :



r_1 tubes avec la dilution d_1

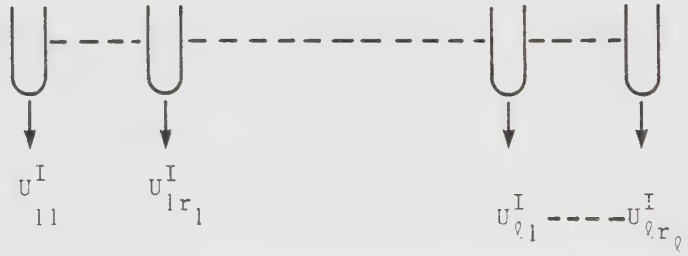
r_ℓ tubes avec la dilution d_ℓ

On ajoute la quantité d'hormone radioactive H^* dans chacun des tubes



la réaction atteint son équilibre.
Séparation de l hormone radioactive libre et du complexe Anticorps-hormone radioactive.

Comptage du complexe Anticorps-hormone radioactive pour chacun des tubes



Estimation du blanc

même schéma que le précédent en remplaçant la préparation inconnue par une solution tampon.

II.8.2. Conséquences sur la variabilité des réponses

Intuitivement nous pouvons imaginer que cette méthode de dosage, dite avec "extraction", introduit des sources d'erreur supplémentaires puisque l'expérimentateur ressent lui-même le besoin de les contrôler en mesurant la récupération et le blanc. Or dans un dosage radioimmunologique, le nombre de préparations inconnues est généralement grand et le nombre de dilutions pour chacune de ces préparations vaut un, rarement deux, jamais plus de deux. Nous ne pouvons donc pas tester l'hypothèse d'égalité des variances résiduelles pour les différentes préparations car le plan d'expérience du dosage radioimmunologique ne nous permet pas d'ajuster une courbe de réponse aux observations des préparations inconnues. Les deux exemples que nous avons traités dans ce chapitre (la testostérone et le cortisol) sont effectués suivant le schéma que nous avons décrit au paragraphe précédent. Lorsque nous avons estimé la variabilité de la variance pour le dosage de la testostérone, nous avons pris soin de vérifier en regroupant plusieurs dosages que les estimations de σ^2 et de h étaient identiques pour les préparations inconnues et les préparations standards. Cela signifie que l'accroissement de la variance est dû pour l'essentiel à l'accroissement de la dose. (Il s'agissait de dosages de la testostérone tous effectués selon la même méthode, dans le même laboratoire, par le même expérimentateur). Toutefois, nous avions constaté (paragraphe II.4.) que l'estimation de σ^2 était légèrement supérieure pour les préparations inconnues que pour les préparations standards. Si nous ne prenons pas en compte cette différence, l'intervalle de confiance que nous calculons autour de la dose inconnue est sûrement faux ; si nous prenons en compte cette différence, nous avons vu au paragraphe II.6. que nous sommes amenés à choisir la méthode proposée par ROSS pour calculer l'intervalle de confiance.

Comme nous le voyons, il serait important de pouvoir répondre convenablement à la question suivante : la variabilité de la variance est-elle identique pour les préparations standard et inconnues d'un même dosage ? Il suffirait pour cela d'effectuer plusieurs dilutions pour une seule des préparations inconnues du dosage. (Un tel plan d'expérience nous permettrait également de tester la condition de similitude : voir annexe B).

II.8.3. Conséquences sur l'estimation des doses inconnues.

Examinons d'abord la pratique courante. Pour la préparation inconnue I la dose, appelée maintenant z^I_{Brute} est estimée par Z^I_{Brute} en inversant la relation dose-réponse estimée pour la préparation standard.

Chaque biologiste, ou presque, possède ensuite sa méthode personnelle pour en déduire une estimation de la dose appelée dose vraie, que nous notons z^I_{Vraie}. Remarquons qu'une telle pratique rend bien difficile toute comparaison des résultats entre différents laboratoires !

Le blanc étant estimé par Z_b, la récupération p_I par $P_I = \frac{v}{v_I} \frac{U^*_I}{\bar{U}^*}$, où $\bar{U}^* = \sum_{i=1}^{m} U^*_i / m$, certains estiment la dose inconnue par $Z^I_{Vraie} = Z^I_{Brute} / P_I$, d'autres par $Z^I_{Vraie} = (Z^I_{Brute} - Z_b)/P_I$. De nombreuses autres versions sont sûrement utilisées, et l'objet du paragraphe n'est pas de déterminer la meilleure d'entre elles. Le principal reproche à faire aux expérimentateurs est de n'avoir jamais tenu compte, lors du calcul ultérieur de la précision estimée pour Z^I_{Vraie}, de la variabilité de P_I (ou de celle de Z_b). Pourtant, il est arrivé à chacun d'entre eux, d'obtenir une valeur calculée de P_I supérieure à 1 ! Les quelques calculs qui suivent tendent à montrer le danger qu'il y a à négliger une telle variabilité.

Calculons un intervalle de confiance pour p_I. On peut supposer que U^*_I est une variable aléatoire de Poisson de paramètre μ^*_I, et nous approcherons la loi de μ^*_I par une loi Normale d'espérance μ^*_I et de variance μ^*_I. (Nous pourrions bien entendu supposer que la variance

de U_I^* est proportionnelle à $\phi_h(\mu_I^*)$; remarquons toutefois que l'observation est ici différente de celle déjà étudiée : la réponse en une dose d'hormone. Il nous faudrait donc refaire une étude longue et fastidieuse pour connaître la forme de ϕ et estimer h. Pour autant, cela n'éclaircirait pas notre problème et l'hypothèse que nous faisons sur la loi de U_I^* est suffisante pour nous permettre d'énoncer les remarques qui vont suivre). De même, nous supposons que pour tout i variant de 1 à m, U_i^* a une loi Normale, d'espérance μ^* et de variance μ^*. Les variables U_I^* et $(U_i^*$, i=1,...,m) sont indépendantes deux à deux. Sous ces hypothèses, on peut alors calculer un intervalle de confiance pour p_I.

Soit $\Delta = \frac{v^2}{v_I^2} \frac{\varphi^2}{4} + \mu^* P_I \left(\frac{v}{v_I} + \frac{P_I}{m} \right)$, soit P_{inf} et P_{sup} définis de la façon suivante :

$$P_{inf} = P_I + \frac{v}{v_I} \frac{\varphi^2}{2\mu^*} - \frac{\varphi}{\mu^*} \sqrt{\Delta}$$

$$P_{sup} = P_I + \frac{v}{v_I} \frac{\varphi^2}{2\mu^*} + \frac{\varphi}{\mu^*} \sqrt{\Delta}$$

où $\varphi = \Phi(\alpha)$.

Alors la probabilité que p_I soit compris entre P_{inf} et P_{sup} vaut $1-\alpha$. Nous remarquons immédiatement que ces intervalles de confiance pour p_I dépendent de la vraie valeur de μ^*. Leur longueur sera d'autant plus faible que μ^* sera grand, que $\frac{v}{v_I}$ sera faible et m grand.

Application numérique : Dosage de la testostérone.

En général, les valeurs observées pour les variables U_i^* sont proches de 500. Nous supposerons donc μ^* connu et égal à 500. Pour ce dosage, m vaut 8 et $\frac{v}{v_I}$ vaut 2. Les valeurs calculées pour p_I varient autour de 0,9.

Valeur calculée pour P_I	Intervalle de confiance calculé, au niveau 95 %.
1	[0,88 ; 1,14]
0,9	[0,79 ; 1,03]
0,8	[0,69 ; 0,92]
0,5	[0,42 ; 0,60]

Pour $\mu^* = 1000$, $m = 8$ et $\frac{y}{v^I} = 2$, l'intervalle de confiance calculé pour une valeur de $P^I = 0,9$ vaut [0,82 ; 0,99]. Le gain sur l'intervalle précédemment calculé n'est pas négligeable. Malheureusement, on ne peut pas augmenter μ^* comme on le désire. En effet, la quantité de radioactivité récupérée après manipulations, doit être suffisamment faible pour ne pas modifier la réponse à une dose d'hormone. Plus précisément cette quantité doit être négligeable devant la quantité d'hormone radioactive notée H^* au paragraphe II.2. Donc si on augmente μ^*, on doit dans les mêmes proportions augmenter le rapport $\frac{y}{v^I}$. Nous avons supposé μ^* connu. Or, on peut calculer aisément un intervalle de confiance pour μ^*.

Soit $\Delta' = \frac{\mathcal{E}^2}{2m}(2\bar{U}^* + \frac{\mathcal{E}^2}{2m})$, soit P'_{inf} et P'_{sup} définis de la façon suivante :

$$P'_{inf} = \bar{U}^* + \frac{\mathcal{E}^2}{2m} - \sqrt{\Delta'} \quad , \quad P'_{sup} = \bar{U}^* + \frac{\mathcal{E}^2}{2m} + \sqrt{\Delta'}$$

où $\mathcal{E} = \mathcal{E}(\alpha)$.

Alors la probabilité que μ^* soit compris entre P'_{inf} et P'_{sup} vaut $1-\alpha$.

Par exemple, si la valeur calculée de $\bar{U}^*$ vaut 500, si m vaut 8 et α vaut 5 %, l'intervalle de confiance calculé pour μ^* est [485 ; 516] .

Supposer dans notre exemple que μ^* est connu et vaut 500 ne peut-être qu'une première approximation.

A partir de quelle valeur de m, l'hypothèse serait-elle raisonnable ? La demi-longueur de l'intervalle de confiance vaut $\sqrt{\Delta'}$. Supposons que nous soyons satisfait avec une précision telle que $\sqrt{\Delta'}$ soit inférieure à $\bar{U}^*/100$; m doit alors rester supérieur à $\xi^2/2\ \bar{U}^*\ (\sqrt{1+10^{-4}}-1)$. Pour notre exemple m doit rester supérieur à 77. Si nous nous contentons d'une précision de l'ordre de 2 % ($\sqrt{\Delta'} \leq \bar{U}^*/50$) m doit être supérieur à 20 et l'intervalle de confiance calculé pour μ^* est [490 ; 510].

Examinons maintenant, comment tenir compte de la variabilité de p_I dans le calcul de l'intervalle de confiance de la dose inconnue. Supposons que le nombre des répétitions m effectuées pour estimer μ^* soit suffisamment grand, pour supposer μ^* connu.

Malgré cette hypothèse, les intervalles de confiance calculés pour p_I sont dans notre exemple trop étendus pour ne pas tenir compte de la variabilité de P_I lors du calcul de la précision de Z^I_{Vraie}.

Supposons que par l'une des deux méthodes proposées au paragraphe II.6. nous ayons calculé un intervalle de confiance pour z^I_{Brute}. Soit Z_{inf} et Z_{sup} tels que la probabilité que z^I_{Brute} soit compris entre Z_{inf} et Z_{sup} vaille $1-\alpha$. Si $Z^I_{Vraie} = Z^I_{Brute}/P_I$ est un estimateur de z^I_{Vraie} , alors z^I_{Vraie} sera compris entre Z_{inf}/P_{sup} et Z_{sup}/P_{inf} avec une probabilité supérieure à $(1-\alpha)^2$.

Reprenons l'exemple du dosage de la testostérone. Supposons que le comptage moyen observé pour la préparation inconnue I soit égal à 650. D'après le paragraphe II.6.2., z^I_{Brute} est estimée par 153 pg par ml, l'intervalle de confiance calculé au niveau 95 % est [140 ; 167].

Supposons que la valeur observée de P_I soit égale à 0,9 (m=8, μ^*=500). Si nous négligeons la variabilité de P_I, nous déduirons que z^I_{Vraie} est estimée par 170 pg par ml, l'intervalle de confiance calculé au niveau 95 % étant [140/0,9 ; 167/0,9] = [155 ; 185] . Si nous prenons en compte la variabilité de P_I, l'intervalle de confiance calculé pour p_I étant égal à [0,79 ; 1,03], on en déduira avec une probabilité supérieure à 0,95 x 0,95 $\simeq$ 0,90 que z^I_{Vraie} appartient à l'intervalle calculé [140/1,03 ; 167/0,79] = [136 ; 211]. L'intervalle de confiance ainsi calculé est très étendu ; la prise en compte de la variabilité de P_I multiplie par 2,5 la longueur de l'intervalle calculé pour notre exemple ! Malheureusement nous ne connaissons pas le niveau de cet intervalle. Nous savons seulement qu'il est supérieur à 90 %, mais il est peut être très proche de 1.

Une autre méthode suggérée par FINNEY (1982) est la suivante : notons X^I_{Vraie} un estimateur du logarithme de z^I_{Vraie} .

$$X^I_{Vraie} = X^I_{Brute} - \text{Log}\,\frac{v}{v_1} - \text{Log}\,U^*_I + \text{Log}\,\bar{U}^*$$

où X^I_{Brute} est un estimateur du logarithme de la dose brute. Le théorème de Fieller permet de calculer un intervalle de confiance pour x^I_{Brute} dont les bornes sont notées X_{inf} et X_{sup}. Nous pouvons alors estimer la variance de X^I_{Brute} par la quantité $(X_{sup}-X_{inf})^2/4\,\tau^2$, où τ a été défini au paragraphe II.6.2. (Si H, défini au paragraphe II.6.2., est négligeable et supposé nul, cette quantité est un estimateur satisfaisant de la variance de X^I_{Brute}).

Ainsi nous pouvons écrire

$$\text{Var}\,X^I_{Vraie} = \text{Var}\,X^I_{Brute} + \text{Var Log}\,U^*_i + \text{Var Log}\,\bar{U}^*$$

Var Log U^*_I étant estimé par $\dfrac{1}{(\ell n 10)^2 U^*_I}$ et Var Log $\bar{U}^*$ par $\dfrac{1}{(\ell n 10)^2 m\,\bar{U}^*}$ (*),

---

(*) *Il suffit de remarquer que la dérivée de Log U vaut 1/U ln10. Aussi ,*

$$\text{Var Log}\,U^*_I \simeq \frac{\text{Var}\,U^*_I}{(\ell n 10)^2 (\mu^*_I)^2} \simeq \frac{1}{(\ell n 10)^2\,\mu^*_I} \quad \text{dont un estimateur est} \quad \frac{1}{(\ell n 10)^2\,U^*_I}\,.$$

la probabilité que x^I_{Vraie} soit compris entre $X^I_{Vraie} - \mathcal{C}\sqrt{\text{Var } X^I_{Vraie}}$ et $X^I_{Vraie} + \mathcal{C}\sqrt{\text{Var } X^I_{Vraie}}$ vaut $1-\alpha$ si $\mathcal{C}$ est défini comme au paragraphe II.6.2.

Pour notre exemple, la valeur calculée pour Var X^I_{Vraie} vaut

$$\frac{(0,077)^2}{4\mathcal{C}^2(0,05;19)} + \frac{1}{(\ell n10)^2\ 225} + \frac{1}{(\ell n10)^2(8x500)}$$

(Nous supposons que la valeur observée pour $\bar{U}^*$ est 500 et la valeur observée pour U^*_I est 225, ainsi la valeur observée pour P_I est 0,9). Ainsi, Var X^I_{Vraie} vaut 0,00122, d'où l'intervalle de confiance calculé pour x^I_{Vraie} : $[1,118 ; 1,265]$. On en déduit un intervalle de confiance pour z^I_{Vraie} : $[131 ; 184]$. Cet intervalle est sensiblement différent de ceux que nous avons déjà calculé.

Nous serions bien incapables de conseiller une méthode plutôt qu'une autre. Nous pensons préférable d'essayer d'obtenir un estimateur de P_I de variance très faible et donc négligeable. Aussi, pour conclure ce paragraphe, nous ferons la remarque suivante : μ^*_I, la quantité d'hormone radioactive présente dans un volume v_I de solution est estimée à l'aide d'une seule observation U^*_I. L'expérience se fait ainsi, car pour chaque préparation inconnue, il faut agiter "à la main" le tube afin que l'hormone collée au verre se dissolve dans le "tampon" (voir paragraphe II.8.1). Mais, il serait possible d'observer plusieurs répétitions de U^*_I, en agitant plusieurs tubes à la fois et en veillant à ce que toutes les manipulations soient effectuées de façon identique pour chacune de ces préparations. μ^*_I serait alors estimé par la moyenne des m_I réponses observées, $\bar{U}^*_I$. On calcule alors un intervalle de confiance pour p_I de la façon suivante :

soit $\Delta = \frac{v^2}{v_I^2}\ \frac{\mathcal{C}^2}{4\ m_I^2} + \mu^*\ p^I\ (\frac{v}{v_I}\ \frac{1}{m_I} + \frac{p^I}{m})$, soit P_{inf} et P_{sup} définis de la façon suivante :

$$P_{inf} = p^I + \frac{v}{v_I} \frac{\xi^2}{2 m_I \mu^*} - \frac{\xi}{\mu^*} \sqrt{\Delta}$$

$$P_{sup} = p^I + \frac{v}{v_I} \frac{\xi^2}{2 m_I \mu^*} + \frac{\xi}{\mu^*} \sqrt{\Delta}$$

où $\xi = \xi(\alpha)$.

Alors, la probabilité que p_I soit compris entre P_{inf} et P_{sup} vaut $1-\alpha$. A partir de quelle valeur de m_I est-il raisonnable de négliger la variabilité de P_I dans le calcul de l'intervalle de confiance pour z^I_{Vraie} ?

Soit ε et ε' définis de la façon suivante : $P_{inf} = p^I - \varepsilon'$, $P_{sup} = p^I + \varepsilon$. Alors avec une probabilité supérieure à $(1-\alpha)^2$,

$$\frac{Z_{inf}}{p^I + \varepsilon} \leq z^I_{Vraie} \leq \frac{Z_{sup}}{p^I - \varepsilon'}$$

ou bien, si ε et ε' sont suffisamment petits

$$\frac{Z_{inf}}{p^I}\left(1 - \frac{\varepsilon}{p^I}\right) \leq z^I_{Vraie} \leq \frac{Z_{sup}}{p^I}\left(1 + \frac{\varepsilon'}{p^I}\right)$$

Ainsi nous pourrons négliger la variabilité de P_I dans le calcul de l'intervalle de confiance pour z^I_{Vraie} si $\frac{\varepsilon}{p^I}$ et $\frac{\varepsilon'}{p^I}$ sont négligeables devant 1.

Il est facile de remarquer que ε est une fonction décroissante de m_I. Ainsi ε est toujours supérieur à $\frac{\xi p^I}{\sqrt{m\mu^*}}$. Si par exemple , $\frac{\varepsilon}{p^I}$ doit rester inférieur à 2 %, $m\mu^*$ doit être supérieur à $\left(\frac{100\,\xi}{2}\right)^2$ qui vaut 10000 si α vaut 5 %.

Dans notre exemple, où la valeur calculée de $\bar{U}^*$ vaut 500, m, le nombre de répétitions effectués pour estimer la quantité d'hormone h^* introduite dans chaque préparation inconnue, doit être supérieur à 20. Si on exige que $\frac{\varepsilon}{p^I}$ reste inférieur à 1 %, m doit être supérieur à 80. Le tableau ci-après donne les valeurs de

ε/p^I et ε'/p^I pour différentes valeurs de m et m_I.

m		m_I = 1	10	20	50	∞
8	$\frac{\varepsilon}{p^I}$	15 %	5,4 %	4,4 %	3,7 %	3,2 %
	$\frac{\varepsilon'}{p^I}$	13 %	5,2 %	4,3 %	3,7 %	3,2 %
20	$\frac{\varepsilon}{p^I}$	14 %	4,8 %	3,6 %	2,8 %	2 %
	$\frac{\varepsilon'}{p^I}$	13 %	4,6 %	3,5 %	4,7 %	2 %
80	$\frac{\varepsilon}{p^I}$	14 %	4,4 %	3,2 %	2,1 %	1 %
	$\frac{\varepsilon'}{p^I}$	12 %	4,2 %	3,1 %	2,1 %	1 %

Les résultats sont calculés, pour une valeur observée de P^I égale à 90 %. (Plus P^I est proche de zéro et plus ε/P_I est grand). Il semblerait, en regardant ce tableau que des valeurs de m et m_I voisines de 20 seraient satisfaisantes pour le cas particulier que nous traitons.

Plus généralement, m étant fixé, comment choisir m_I pour que le rapport $\frac{\varepsilon}{p^I}$ reste inférieur à x/100 ?

Les calculs montrent que x étant fixé, le problème n'a de solution que si $m\,\mu^*$ est supérieur à $\varphi^2\ 10^4/x^2$. On a alors $\varepsilon/p^I \leq x/100$ si

$$m_I \geq \frac{\varphi^2}{\mu^*}\ \frac{v}{v_I}\ (1 + \frac{x}{10^2})\ /\ p^I\ (\frac{x^2}{10^4} - \frac{\varphi^2}{m\,\mu^*})$$

si x = 2 , m = 25 , alors $m_I \geq 230$

si x = 4 , m = 10 , alors $m_I \geq 24$

Il semblerait donc que la voie la plus raisonnable est d'essayer d'obtenir des répétitions de la mesure de U_I^*. Ceci est réalisable s'il est possible de manipuler dans les mêmes conditions exactement plusieurs préparations inconnues. La récupération sera alors estimée non pas pour une préparation inconnue mais pour plusieurs préparations inconnues et il sera alors possible de négliger la variabilité de cet estimateur lors du calcul de l'intervalle de confiance pour z^I_{Vraie}.

A N N E X E S

-=-=-=-

ANNEXE A

Etude de la variabilité de la variance

Supposons que pour une dose z_i d'hormone (connue ou non), nous notions U_{ij} la jème réponse observée, j varie de 1 à k_i, i de 1 à k.

On suppose U_{ij} Normalement distribuée, d'espérance μ_i et de variance proportionnelle à $\phi_h(\mu_i)$

$$U_{ij} \sim \mathcal{N}(\mu_i, \sigma^2 \phi_h(\mu_i))$$

$$\phi_h(\mu_i) = \mu_i^h \quad \text{où} \quad \phi_h(\mu_i) = \mu_i(1 + h\,\mu_i)$$

i étant fixé, la densité de probabilité des $(U_{i1},\ldots,U_{ik_i})$ est

$$\mathcal{V}_i = (2\pi\sigma^2 \phi_h(\mu_i))^{-k_i/2} \exp-(S_i/2\sigma^2 \phi_h(\mu_i))$$

$$\text{où} \qquad S_i = \sum_{j=1}^{k_i} (U_{ij} - \mu_i)^2$$

Soit $\mathcal{L}$ le logarithme de la vraisemblance, $\mathcal{L} = \sum_{i=1}^{k} \mathcal{L}_i$ où

$$\mathcal{L}_i = -\frac{k_i}{2} \ln \sigma^2 - \frac{k_i}{2} \ln \phi_h(\mu_i) - S_i/2\sigma^2 \phi_h(\mu_i) \quad \text{à une constante près.}$$

Soit Σ^2 et H les estimateurs de σ^2 et h qui maximisent la vraisemblance $\mathcal{L}$. RAAB (1981) montre par simulations que si tous les k_i sont égaux, H est un estimateur sans biais de h mais Σ^2 est biaisé d'un facteur $(k_i-1)/k_i$, si tous les k_i ne sont pas tous égaux, H et Σ^2 sont des estimateurs biaisés de h et σ^2.

Considérons la densité de probabilité modifiée, $\mathcal{V}'_i$

$$\mathcal{V}'_i = (2\pi\ \sigma^2\ \phi_h(\mu_i))^{-(k_i-1)/2} \quad \exp - (S_i/2\sigma^2\ \phi_h(\mu_i))$$

Soit $\mathcal{L}'$ le logarithme de la vraisemblance modifiée ,

$$\mathcal{L}' = \sum_i \mathcal{L}'_i \quad \text{où } \mathcal{L}'_i = -\frac{k_i-1}{2} \ln \sigma^2 - \frac{k_i-1}{2} \ln \phi_h(\mu_i) - S_i/2\ \sigma^2 \phi_h(\mu_i)$$

à une constante près. Soit $\Sigma^{2'}$ et H' les estimateurs du maximum de vraisemblance modifiée de σ^2 et h. Par simulation, on montre que quelle que soit la valeur des k_i, $\Sigma^{2'}$ et H' sont des estimateurs approximativement sans biais de σ^2 et h pour des grandes valeurs de k, de plus, la matrice de variance-covariance de H et Σ^2 est approximée par

$$-E\begin{pmatrix} \dfrac{\partial^2 \mathcal{L}'}{\partial h^2} & \dfrac{\partial^2 \mathcal{L}'}{\partial h \partial \sigma^2} \\ \dfrac{\partial^2 \mathcal{L}'}{\partial h \partial \sigma^2} & \dfrac{\partial^2 \mathcal{L}'}{\partial (\sigma^2)^2} \end{pmatrix}^{-1}$$

Appelons $\mathcal{H}$ l'hypothèse '$U_{ij} \sim \mathcal{N}(\mu_i,\ \sigma^2\ \phi_h(\mu_i))$' et $\mathcal{A}$ l'hypothèse '$U_{ij} \sim \mathcal{N}(\mu_i, \sigma_i^2)$'. RAAB présente également dans son article un test asymptotique de $\mathcal{H}$ contre $\mathcal{A}$. Soit $\mathcal{V}_{Ai}'$ la densité de probabilité modifiée de $(U_{i1},\ldots U_{ik_i})$ sous l'hypothèse $\mathcal{A}$

$$\mathcal{V}_{Ai}' = (2\pi\ \sigma_i^2)^{-(k_i-1)/2} \quad \exp - s_i/2\sigma_i^2$$

où $\quad s_i = \sum_{j=1}^{k_i} (U_{ij}-U_{i.})^2$

Soit $\mathcal{L}'_A$ le logarithme de la vraisemblance modifiée sous l'hypothèse $\mathcal{H}$, $\mathcal{L}'_A = \sum_{i=1}^{k} \mathcal{L}'_{A_i}$ où $\mathcal{L}'_{A_i} = -\frac{k_i - 1}{2} \log \sigma_i^2 - s_i/2\,\sigma_i^2$ à une constante près.

Soit $\Sigma_i^{2'}$ l'estimateur du maximum de vraisemblance modifiée de σ_i^2, $\Sigma_i^{2'} = s_i/(k_i - 1)$

Appelons $\tilde{\mathcal{L}}'$ (respectivement $\tilde{\mathcal{L}}'_A$) la valeur de la vraisemblance modifiée en $\Sigma^{2'}$, H', $U'_{i.}$, $i=1,\ldots k$ (respectivement en $\Sigma_i^{2'}$) , et Γ la statistique $-2(\tilde{\mathcal{L}}' - \tilde{\mathcal{L}}'_A)$.

Quand k est grand, la loi de Γ peut être approximée par une loi Normale dont l'espérance $E(\Gamma)$ et la variance $V(\Gamma)$ sont tabulées. On rejettera par exemple l'hypothèse $\mathcal{H}$ au niveau 5 % quand $|\Gamma - E(\Gamma)|$ sera supérieur à $2\sqrt{\mathrm{Var}(\Gamma)}$.

ANNEXE B

Estimation des paramètres de la courbe standard et des préparations inconnues.

1er cas : On observe les réponses pour la préparation standard et une préparation inconnue.

Soit U^s_{ij} les réponses de la préparation standard, en k doses z^s_i , j variant de 1 à n_i,

U^I_{ij} les réponses de la préparation inconnue en ℓ dilutions d_i, j variant de 1 à r_i.

On suppose U^s_{ij} et U^I_{ij} normalement distribuées.

$E(U^s_{ij}) = \mu^s_i$, $E(U^I_{ij}) = \mu^I_i$. Appelons F la fonction décrivant la relation entre la dose et la réponse

$$\mu^s = F^s(z) \quad , \quad \mu^I = F^I(z)$$

Si les conditions de validité du dosage sont vérifiées, il existe un nombre ρ tel que pour tout z $F^I(z) = F^s(\rho z)$; ρ est l'efficacité de la préparation inconnue relativement à la préparation standard, c'est-à-dire la quantité de préparation standard équivalente en effet à une unité de préparation inconnue. Une estimation de ρ nous fournira donc une estimation de la dose inconnue.

Si $\mu^s = c + (d-c)/(1+\exp-(\alpha^s+ \beta x))$, $x = \text{Log}(z)$, alors les conditions de validité du dosage étant vérifiées

$$\mu^I = c + (d-c)/(1+\exp- (\alpha^I+\beta \text{ Log } d))$$

$\alpha^I = \alpha^s + \beta \text{ Log}(\rho)$. Ainsi $\text{Log}(\rho) = (\alpha^I-\alpha^s)/\beta$

Supposons de plus que la variabilité des réponses est la même pour la préparation standard et pour la préparation inconnue : $\text{Var}(U^s_{ij}) = \sigma^2 \phi_h(\mu^s_i)$ et $\text{Var}(U^I_{ij}) = \sigma^2 \phi_h(\mu^I_i)$.

Le modèle étant ainsi posé, nous allons à l'aide des réponses observées pour la préparation standard et pour la préparation inconnue, estimer les six paramètres C,D,α^I,α^s, β et $\sigma 2$. Supposant les réponses indépendamment observées, l'opposé du logarithme de la vraisemblance $\mathcal{L} = \mathcal{L}^s + \mathcal{L}^I$

où

$$2\mathcal{L}^s = \sum_{i=1}^{k} n_i \log 2\pi\sigma^2 + \sum_{i=1}^{k} n_i \ln \phi_h(\mu_i^s) + \sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{(U_{ij}^s - \mu_i^s)^2}{\sigma^2 \phi_h(\mu_i^s)}$$

$$2\mathcal{L}^I = \sum_{i=1}^{\ell} r_i \log 2\pi\sigma^2 + \sum_{i=1}^{\ell} r_i \ln \phi_h(\mu_i^I) + \sum_{i=1}^{\ell}\sum_{j=1}^{r_i} \frac{(U_{ij}^I - \mu_i^I)^2}{\sigma^2 \phi_h(\mu_i^I)}$$

Le problème est maintenant de chercher C,D,A^I,B,A^s et S^2 qui minimisent $\mathcal{L}$. Il n'est pas possible d'exprimer analytiquement ces estimateurs en fonction des réponses U_{ij}, nous devons donc utiliser un programme d'optimisation dont il n'est pas opportun ici de décrire la structure. Appelons R l'estimateur du logarithme de ρ , $R = (A^I - A^s)/B$. Pour estimer un intervalle de confiance pour $\ln \rho$, nous appliquons l'une des deux méthodes proposée au paragraphe II.6. au paramètre $r = (\alpha^I - \alpha^s)/\beta$. Nous devons donc estimer les variances et covariances des estimateurs. La théorie générale du modèle non linéaire nous permet d'estimer la matrice de variance-covariance Σ des estimateurs par la matrice symétrique définie positive.

$$\Sigma = + \left[\begin{array}{cccccc} \frac{\partial^2 \mathcal{L}}{\partial c^2} & & & & & \\ \frac{\partial^2 \mathcal{L}}{\partial c \partial Bo} & \frac{\partial^2 \mathcal{L}}{\partial d^2} & & & & \\ \frac{\partial^2 \mathcal{L}}{\partial c \partial \alpha^s} & \frac{\partial^2 \mathcal{L}}{\partial d \partial \alpha^s} & \frac{\partial^2 \mathcal{L}}{\partial \alpha^{s2}} & & & \\ \frac{\partial^2 \mathcal{L}}{\partial c \partial \alpha^I} & \frac{\partial^2 \mathcal{L}}{\partial d \partial \alpha^I} & \frac{\partial^2 \mathcal{L}}{\partial \alpha^s \partial \alpha^I} & \frac{\partial^2 \mathcal{L}}{\partial \alpha^{I2}} & & \\ \frac{\partial^2 \mathcal{L}}{\partial c \partial \beta} & \frac{\partial^2 \mathcal{L}}{\partial d \partial \beta} & \frac{\partial^2 \mathcal{L}}{\partial \alpha^s \partial \beta} & \frac{\partial^2 \mathcal{L}}{\partial \alpha^I \partial \beta} & \frac{\partial^2 \mathcal{L}}{\partial \beta^2} & \\ \frac{\partial^2 \mathcal{L}}{\partial c \, \partial \sigma^2} & \frac{\partial^2 \mathcal{L}}{\partial d \partial \sigma^2} & \frac{\partial^2 \mathcal{L}}{\partial \alpha^s \partial \sigma^2} & \frac{\partial^2 \mathcal{L}}{\partial \alpha^I \partial \sigma^2} & \frac{\partial^2 \mathcal{L}}{\partial \beta \partial \sigma^2} & \frac{\partial^2 \mathcal{L}}{\partial (\sigma^2)^2} \end{array} \right]^{-1}$$

où les dérivées secondes de $\mathcal{L}$ sont calculées au point (C,D,A^s,A^I,B,S^2).

Une autre méthode d'estimation des paramètres proposée par FINNEY (1977), consiste à supposer que minimiser $\mathcal{L}$ en $c,d,\alpha^s,\alpha^I,\beta$ est équivalent à minimiser $\mathcal{Y} = \mathcal{Y}^s + \mathcal{Y}^I$

où $$\mathcal{Y}^s = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \frac{(U_{ij}^s - \mu_i^s)^2}{\sigma^2 \, \phi_h(\mu_i^s)} \quad \text{et} \quad \mathcal{Y}^I = \sum_{i=1}^{\ell} \sum_{j=1}^{r_i} \frac{(U_{ij}^I - \mu_i^I)^2}{\sigma^2 \, \phi_h(\mu_i^I)}$$

FINNEY assure que ceci est vrai si le nombre d'observations est grand, c'est-à-dire si $\sum_{i=1}^{k} n_i$ et $\sum_{i=1}^{\ell} r_i$ sont grands. $\frac{1}{\sigma^2}$ se mettant alors en facteur, on minimise par rapport à c,d,α^s,α^I et β, soit C,D,As,A^I et B les estimateurs,

σ^2 est alors estimé par :

$$S^2 = \left(\sum_{i,j} \frac{(U^s_{ij} - \hat{\mu}^s_i)^2}{\phi_h(\hat{\mu}^s_i)} + \sum_{i,j} \frac{(U^I_{ij} - \hat{\mu}^I_i)^2}{\phi_h(\hat{\mu}^I_i)} \right) / \left(\sum_{i=1}^{k} n_i + \sum_{i=1}^{\ell} r_i - 5 \right)$$

où $\hat{\mu}^s_i$ (respectivement $\hat{\mu}^I_i$) est l'estimateur de μ^s_i (respectivement μ^I_i) c'est-à-dire, la valeur de U^s_i (respectivement U^I_i) calculée en C,D,A^s,B (respectivement C,D,A^I,B).

L'estimation de c,d,α^s,α^I, et β étant indépendante de σ^2, l'estimation de la matrice de variance-covariance des estimateurs C,D,A^s,A^I et B est Σ

$$\Sigma = \begin{bmatrix} \frac{\partial^2 \mathcal{L}}{\partial c^2} & & & & \\ \frac{\partial^2 \mathcal{L}}{\partial c \partial Bo} & \frac{\partial^2 \mathcal{L}}{\partial d^2} & & & \\ \frac{\partial^2 \mathcal{L}}{\partial c \partial \alpha^s} & \frac{\partial^2 \mathcal{L}}{\partial d \partial \alpha^s} & \frac{\partial^2 \mathcal{L}}{\partial \alpha^{s2}} & & \\ \frac{\partial^2 \mathcal{L}}{\partial c \partial \alpha^I} & \frac{\partial^2 \mathcal{L}}{\partial d \partial \alpha^I} & \frac{\partial^2 \mathcal{L}}{\partial \alpha^s \partial \alpha^I} & \frac{\partial^2 \mathcal{L}}{\partial \alpha^{I2}} & \\ \frac{\partial^2 \mathcal{L}}{\partial c \partial \beta} & \frac{\partial^2 \mathcal{L}}{\partial d \partial \beta} & \frac{\partial^2 \mathcal{L}}{\partial \alpha^s \partial \beta} & \frac{\partial^2 \mathcal{L}}{\partial \alpha^I \partial \beta} & \frac{\partial^2 \mathcal{L}}{\partial \beta^2} \end{bmatrix}^{-1}$$

Remarques :

1) Les observations obtenues pour la préparation inconnue contribuent à l'estimation de c,d et β. Il est donc nécessaire que le nombre ℓ de dilutions d_i soit suffisamment grand pour que les réponses observées pour la préparation inconnue s'échelonnent entre c et d.

2) Un tel plan d'expérience permet de vérifier statistiquement les conditions de validité du dosage, c'est-à-dire le "parallélisme" des courbes de réponses.

Soit l'hypothèse $\mathcal{H}$ "les courbes sont parallèles" contre l'alternative $\mathcal{A}$ "les courbes ne sont pas parallèles" (il serait plus rigoureux ici d'utiliser le terme "translatées". Pour plus de clarté, les hypothèses $\mathcal{H}$ et $\mathcal{A}$ sont explicitées ci-dessus).

Sous l'hypothèse $\mathcal{H}$,

$$\mu^S = c^S + (d^S - c^S)/(1+\exp-(\alpha^S + \beta^S x))$$

$$\mu^I = c^I + (d^I - c^I)/(1+\exp-(\alpha^I + \beta^I d))$$

L'hypothèse $\mathcal{H}$ se traduit par "$(c^S, d^S, \beta^S) = (c^I, d^I, \beta^I) = (c, d, \beta)$"

et $\mathcal{A}$ par "$(c^S, d^S, \beta^S) \neq (c^I, d^I, \beta^I)$" . Supposons Σn_i et Σr_i grands.

Soit $\mathcal{S}_H = (\sum_{i=1}^{k} n_i + \sum_{i=1}^{\ell} r_i - 5)\ S^2$

Soit $\tilde{C}$, $\tilde{D}$, $\tilde{A}^S$ et $\tilde{B}^S$ les estimateurs de c^S, d^S, α^S et β^S qui minimisent

$\mathcal{S}^S$ et $\tilde{C}^I$, $\tilde{D}^I$, $\tilde{A}^I$ et $\tilde{B}^I$ les estimateurs de c^I, d^I, α^I et β^I qui minimisent $\mathcal{S}^I$.

Notons S^S et S^I les sommes des carrés résiduelles

$$S^S = \sum_i \sum_j \frac{(U^S_{ij} - \tilde{\mu}^S_i)^2}{\phi_h(\tilde{\mu}^S_i)} \quad , \quad \tilde{\mu}^S_i = \tilde{C}^S + (\tilde{D}^S - \tilde{C}^S)/(1+\exp-(\tilde{A}^S + \tilde{B}^S x_i))$$

$$S^I = \sum_i \sum_j \frac{(U^I_{ij} - \tilde{\mu}^I_i)^2}{\phi_h(\tilde{\mu}^I_i)} \quad , \quad \tilde{\mu}^I_i = \tilde{C}^I + (\tilde{D}^I - \tilde{C}^I)/(1+\exp-(\tilde{A}^I + \tilde{B}^I d_i))$$

Soit $\mathcal{S}_A = (S^S + S^I)$

La statistique $\mathcal{U} = \dfrac{(\mathcal{S}_H - \mathcal{S}_A)/3}{\mathcal{S}_A/(\Sigma n_i + \Sigma r_i - 8)}$ doit être comparée à la valeur

d'un Fisher à 3 et $(\Sigma n_i+\Sigma r_i-8)$ degrés de liberté.

3) Un tel plan d'expérience n'est jamais construit dans la pratique. C'est dommage, il serait pourtant rassurant dans certains cas de pouvoir vérifier les conditions de validité du dosage, surtout dans les cas où la méthode de dosage pour les préparations inconnues est différente de celle de la préparation standard. De plus, il ne serait pas très coûteux pour un dosage où l'on estime une centaine de doses inconnues, d'effectuer une dizaine de dilutions pour une préparation inconnue.

2ème cas : Pour quelques préparations inconnues, la réponse est mesurée en deux ou trois dilutions.

Ce plan d'expérience est celui qui est habituellement employé. Pour la grande majorité des doses inconnues ℓ est égal à 1 pour des raisons d'économie de temps. Parfois ℓ vaut 2 ou 3. Etudions d'abord le cas où ℓ est supérieur à 2.

Soit U^I_{ij} les réponses pour les ℓ dilutions d_i, j variant de 1 à r_i. En général, l'expérimentateur voulant estimer la dose de la préparation I avec une bonne précision, retiendra pour comptage observé la valeur de $U^I_{i.}$ la plus proche de l'ordonnée du milieu des deux asymptotes c et d. Cette démarche est raisonnable et évite souvent de recommencer le dosage d'une préparation inconnue pour laquelle le volume de dosage a été mal choisi. Contrairement à l'apparence, elle n'est pas dépensière. Toutefois en ne retenant pas toutes les observations, nous perdons de l'information, donc nous diminuons sûrement la précision de notre estimation. D'autre part, la valeur de ℓ est trop faible pour appliquer la méthode précédemment exposée. Aussi nous allons estimer les paramètres de la courbe standard avec les seules observations de la préparation standard.

Soit U^S_{ij} les réponses de la préparation standard, en k doses z^S_i, j variant de 1 à n_i. On suppose U^S_{ij} Normalement distribuée, $E(U^S_{ij}) = \mu^S_i$.

$$\mu^S_i = c + (d-c)\ (1+\exp - (\alpha+\beta\ x_i) \quad \text{où } x_i = \text{Log}\ (z_i)$$

Supposons de plus $\mathrm{Var}(U^s_{ij}) = \sigma^2 \phi_h(\mu^s_i)$

Les estimateurs du maximum de vraisemblance C,D,A,B et S^2 de $c,d,\alpha,\beta,\sigma^2$ minimisent $\mathcal{L}^s$, et l'estimation Σ de la matrice de variance-covariance de ces estimations est donnée par

$$\Sigma = \left[\begin{array}{ccccc} \dfrac{\partial^2 \mathcal{L}^s}{\partial c^2} & & & & \\ \dfrac{\partial^2 \mathcal{L}^s}{\partial c \partial B_o} & \dfrac{\partial^2 \mathcal{L}^s}{\partial d^2} & & & \\ \dfrac{\partial^2 \mathcal{L}^s}{\partial c \partial \alpha} & \dfrac{\partial^2 \mathcal{L}^s}{\partial d \partial \alpha} & \dfrac{\partial^2 \mathcal{L}^s}{\partial \alpha^2} & & \\ \dfrac{\partial^2 \mathcal{L}^s}{\partial c \partial \beta} & \dfrac{\partial^2 \mathcal{L}^s}{\partial d \partial \beta} & \dfrac{\partial^2 \mathcal{L}^s}{\partial \alpha \partial \beta} & \dfrac{\partial^2 \mathcal{L}^s}{\partial \beta^2} & \\ \dfrac{\partial^2 \mathcal{L}^s}{\partial c \partial \sigma^2} & \dfrac{\partial^2 \mathcal{L}^s}{\partial d \partial \sigma^2} & \dfrac{\partial^2 \mathcal{L}^s}{\partial \alpha \partial \sigma^2} & \dfrac{\partial^2 \mathcal{L}^s}{\partial \beta \partial \sigma^2} & \dfrac{\partial^2 \mathcal{L}^s}{\partial (\sigma^2)^2} \end{array} \right]^{-1}$$

Soit μ^I_i l'espérance de U^I_{ij}. Supposant les conditions de validité du dosage vérifiées,

$$\mu^I_i = c + (d-c)(1+\exp-(\alpha^I+\beta\, d_i))$$

Ne disposant pas d'un nombre suffisant d'observations pour estimer les 5 paramètres c,d,α^I,β et σ^2, nous allons remplacer dans l'expression de μ^I_i les paramètres c,d et β par leurs estimateurs C,D et B obtenus à l'aide des observations de la préparation standard. Ainsi l'estimateur du maximum de vraisemblance

A^I de α^I minimise

$$\bar{\mathcal{L}}^I = \frac{1}{2} \sum_{i=1}^{\ell} r_i \log \phi_h(\bar{\mu}_i^I) + \frac{1}{2} \sum_i \sum_j \frac{(U_{ij}^I - \bar{\mu}_i^I)^2}{s^2 \phi_h(\bar{\mu}_i^I)}$$

où $\bar{\mu}_i^I = C + (D-C)\ (1+\exp-(\alpha^I + B\ d_i))$

l'estimateur du logarithme de ρ est alors égal à

$$(A^I - A)/B \quad \text{et} \quad \mathrm{Var}(A^I) = \frac{\partial^2 \bar{\mathcal{L}}^I}{\partial A^{I2}}$$

Si $\sum_{i=1}^{k} n_i$ est grand, minimiser $\mathcal{L}^s$ en c,d,α,β est équivalent à minimiser $\mathcal{S}^s$.

Alors, l'estimation de c,d,α,β est indépendante de σ^2, et

$$s^2 = \sum_{i,j} \frac{(U_{ij} - \hat{\mu}_i^s)^2}{\phi_h(\hat{\mu}_i^s)} \Big/ \left(\sum_{i=1}^{k} n_i - 4 \right)$$

où $\hat{\mu}_i^s = C + (D-C)(1+\exp - (A+B\ x_i))$

C, D, A et B minimisant $\sum_{i,j} \frac{(U_{ij} - \mu_i^s)^2}{\phi_h(\mu_i^s)}$, et la matrice de variance-covariance

des estimateurs de c,d,α,β est estimée par

$$\Sigma = \begin{bmatrix} \frac{\partial^2 \mathcal{L}^s}{\partial c^2} & & & \\ \frac{\partial^2 \mathcal{L}^s}{\partial c \partial B_o} & \frac{\partial^2 \mathcal{L}^s}{\partial d^2} & & \\ \frac{\partial^2 \mathcal{L}^s}{\partial c \partial \alpha} & \frac{\partial^2 \mathcal{L}^s}{\partial d \partial \alpha} & \frac{\partial^2 \mathcal{L}^s}{\partial \alpha^2} & \\ \frac{\partial^2 \mathcal{L}^s}{\partial c \partial \beta} & \frac{\partial^2 \mathcal{L}^s}{\partial d \partial \beta} & \frac{\partial^2 \mathcal{L}^s}{\partial \alpha \partial \beta} & \frac{\partial^2 \mathcal{L}^s}{\partial \beta^2} \end{bmatrix}^{-1}$$

Le cas où ℓ vaut 1 a déjà été largement étudié tout au long de ce chapitre.

Un programme conversationnel, DOSSTAT, écrit en Fortran est disponible. L'annexe suivante en donne la description.

A N N E X E C

LE PROGRAMME DOSSTAT

Ce programme écrit en FORTRAN, est actuellement utilisable sur IRIS 80. Prochainement, une version Mini 6 et une version Multics seront créees.

L'utilisateur doit appeler la commande DOSSTAT, et fournir le nom du fichier sur lequel sont stockées les données. Actuellement ce fichier doit être créé à l'aide du programme DOSDON. DOSDON, est un programme conversationnel, écrit en Fortran, utilisable par de non-informaticiens.

Une seconde version de DOSSTAT, appelée DOSSTAT2 existe. Elle diffère essentiellement de la première par la méthode de calcul des intervalles de confiance pour les doses inconnues.

Que fait DOSSTAT ? Nous distinguerons trois étapes :

ETAPE 1

. Lecture et impression des données

. notice d'utilisation du programme sur demande

ETAPE 2

Etude de la préparation standard

ETAPE 2.1.

Choix du modèle

ETAPE 2.2.

Estimation des paramètres

ETAPE 2.3.

Test d'adéquation au modèle

ETAPE 3

Estimation des doses inconnues

FIN DU PROGRAMME

ETAPE I

Impression des données de la préparation standard

Soit U_{ij} la jème répétition de la réponse pour une dose z_i. j varie de 1 à n_i pour i variant de 1 à k.

Contraintes modifiables :

$k \leq 30$; $n_i \leq 16$ pour tout i ;

$\sum_{i=1,k} n_i \leq 100$; $1 < U_{ij} < 10\,000$;

$0 \leq z_i < 20\,000$

contrainte non modifiable : il est obligatoire de mesurer la réponse pour une dose nulle d'hormone.

Impression des données des préparations inconnues

Soit U^I_{ij} la jème répétition de la réponse pour la dilution d^I_i de la préparation inconnue I. j varie de 1 à r^I_i, i varie de 1 à ℓ^I pour I variant de 1 à M.

contraintes modifiables :

$M \leq 200$; $\ell^I \leq 5$ pour tout I ;

$r^I_i \leq 8$ pour tout i ;

Si l'utilisateur le demande, le programme imprime la notice d'utilisation

↓

ETAPE 2

ETAPE 2 : Etude de la préparation standard

Choix du modèle

Etape 2.1. : différentes possibilités

$E(U/z) = \mu$

Etape 2.1.1. Variabilité de la variance

- La variance est constante — oui → $Var(U/z) = \sigma^2$
- non → La variance est proportionnelle à μ^h — oui → $Var(U/z) = \sigma^2 \mu^h$
- non → La variance est proportionnelle à $\mu(1+h\mu)$ — oui → $Var(U/z) = \sigma^2 \mu (1+h\mu)$

Etape 2.1.2. : Variations de l'espérance de la réponse en fonction de la dose

$E(U/z) = \mu$ $\quad f(y) = 1/(1+\exp-(\alpha+\beta x))$ $\quad x = \text{Log } z$

- La courbe est symétrique — oui → $g(y) = f(y)$
- non →
  - $g(y) = f(y)^{\gamma}$ $\quad \gamma<0$
  - $g(y)=1-(1-f(y))^{-\gamma}$ $\quad \gamma<0$

- La réponse pour une dose infinie d'hormone est mesurée — non → la réponse est supposée nulle — oui → $\mu = dg(y)$
- La réponse pour une dose infinie d'hormone est mesurée — non → $\mu=c+(d-c)g(y)$

Etape 2.1 : Ce que l'utilisateur doit faire

possibilité n° 1 : option par défaut

Le programme choisit lui-même le modèle de la façon suivante :

$$g(y) = f(y)$$
$$\mu = c + (d-c)\, g(y)$$
$$\mathrm{Var}(U/z) = \sigma^2 \mu$$

possibilité n° 2 : l'utilisateur choisit lui-même le modèle

1. Choix de h

Si l'utilisateur le désire le programme estime la variabilité de la variance selon la méthode exposée dans l'annexe A.
La statistique d'adéquation au modèle, ainsi que le résultat du test sont imprimés.

Quel que soit le résultat, l'utilisateur introduit lui-même la valeur de h

2. Si la réponse pour une dose infinie d'hormone n'est pas mesurée, l'utilisateur indique si c est supposé nul ou pas.

3. Choix de γ

Si l'utilisateur le désire le programme estime γ.

. estimation de h selon la méthode exposée dans l'annexe A
. U est supposée Normalement distribuée, d'espérance μ, $\mu = c + (d-c)g(y)$ et de variance proportionnelle à μ_o^h où μ_o est la moyenne des réponses observées pour une dose z d'hormone. Sous ces hypothèses, estimation de c,d,α,β et γ à l'aide de la méthode de Marquardt (1963).
. Le calcul de l'intervalle de confiance pour γ permet à l'utilisateur de décider si la courbe est symétrique ou non.

Quel que soit le résultat, l'utilisateur introduit lui-même la valeur de γ.

Etape 2.2. : estimation des paramètres

Le modèle ayant été choisi, le programme minimise la somme des carrés $\sum_{i,j} \frac{(U_{ij} - \mu_i)^2}{\phi_h(\mu_i)}$ par rapport aux paramètres d,α,β et éventuellement c. (Utilisation du sous-programme VA05AD de la programmathèque HARWELL, 1981). L'estimation de la matrice de variance-covariance des estimateurs D,A,B et éventuellement C est calculée en inversant la matrice de dérivées secondes du logarithme de la vraisemblance (voir annexe B)

σ^2 est estimé par $S^2 = \frac{1}{n-p} \sum_{i,j} \frac{(U_{ij} - \hat{\mu}_i)^2}{\phi_h(\hat{\mu}_i)}$

où $\hat{\mu}_i = C + (D-C)\, g(A + B\, x_i)$

ou bien $\hat{\mu}_i = D\, g(A + B\, x_i)$

$n = \sum_{i=1,k} n_i$ est le nombre total des observations pour la préparation standard

p est le nombre des paramètres (p vaut 3 ou 4)

Le programme imprime les estimations des paramètres ainsi que leur écart-type, puis la valeur calculée de S^2

Etape 2.3. : Test d'adéquation au modèle

Le programme calcule les statistiques de test d'adéquation au modèle $\mathcal{S}_H$ et $\mathcal{S}_A$ (voir paragraphe II.5.2.).

Il imprime les valeurs de ces statistiques, ainsi que le résultat du test ; puis le tableau des résidus réduits et le graphe de la courbe standard.

L'utilisateur a alors deux possibilités :

Estimer les doses inconnues s'il y a lieu → ETAPE 3

Estimer à nouveau les paramètres de la courbe standard après avoir choisi un nouveau modèle. → Possibilité n° 2 de l'étape 2.1.

ETAPE 3 : Estimation des doses inconnues

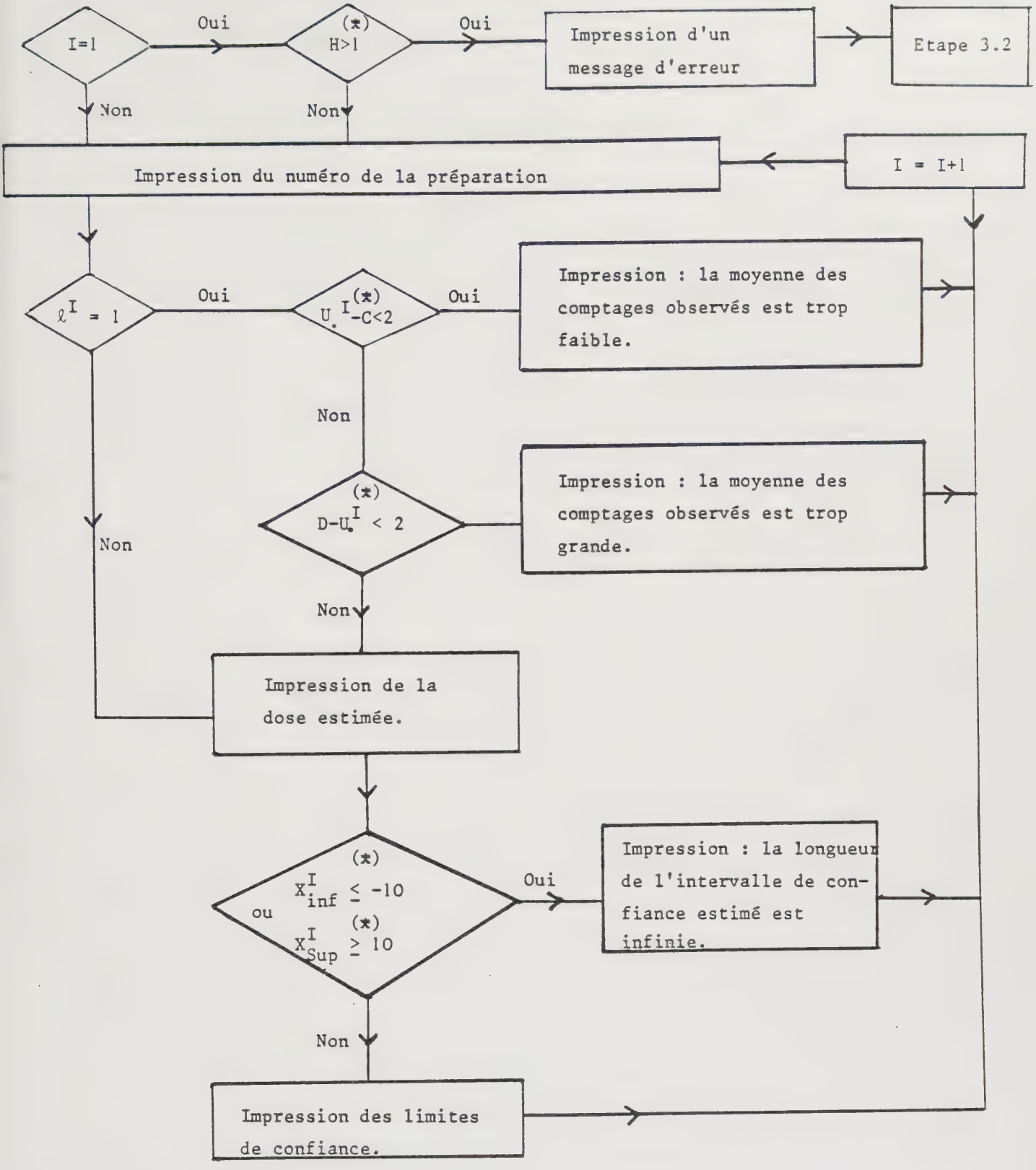
Méthodes utilisées

1er cas : $\ell^I=1$

Estimation de la dose inconnue (paragraphe II.6.1.) calcul d'un intervalle de confiance à l'aide du théorème de Fieller (paragraphe II.6.2.

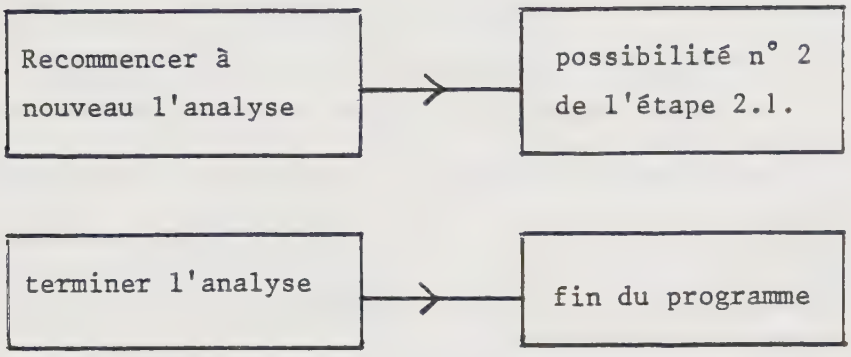
2ème cas : $\ell^I>1$

Le programme estime le paramètre α^I (voir 2ème cas de l'annexe B) et en déduit un estimateur du logarithme de la dose $X^I=(A^I-A)/B$. Le théorème de Fieller appliqué au rapport de A^I-A sur B fournit un intervalle de confiance pour X^I, donc pour Z^I.



(*) *Si H est supérieur à 1, cela signifie que B n'est pas significativement différent de 0 (voir paragraphe 2.6.2.).*
U^I est le moyenne des comptages observés pour la préparation inconnue I.
X^I_{inf} et X^I_{sup} sont les limites inférieures et supérieures du logarithme de la dose inconnue. Nous avons vu que $Z^I_{inf} = 10^{X^I_{inf}}$, $Z^I_{sup} = 10^{X^I_{sup}}$ sont les limites pour la dose inconnue que l'on calculera pour les valeurs de $X^I_{inf} > -10$ ou de $X^I_{sup} < 10$.

L'utilisateur a alors deux possibilités :



Que fait DOSSTAT2 ? Nous distinguerons trois étapes :

ETAPE 1

. Lecture et impression des données

. notice d'utilisation du programme sur demande

ETAPE 2

Etude de la préparation standard

ETAPE 2.1.

Choix du modèle

ETAPE 2.2.

Estimation des paramètres

ETAPE 2.3.

Test d'adéquation au modèle

ETAPE 2.4.

Etude des limites d'utilisation de la courbe standard

ETAPE 3

Estimation des doses inconnues

FIN DU PROGRAMME

ETAPE 1

ETAPE 2

ETAPE 2.1.
ETAPE 2.2.
ETAPE 2.3.

} même chose que pour DOSSTAT

Etape 2.4. : Etude des limites d'utilisation de la courbe standard.

A l'aide de la méthode exposée dans le paragraphe II.6.3., le programme calcule la plus faible dose significativement différente de zéro à 95 %, et la plus forte dose significativement différente de la dose infinie (le calcul est limité à l'intervalle des doses choisi par l'expérimentateur pour estimer la courbe standard).

Pour une réponse moyenne variant de 5 % à 95 % (il s'agit du pourcentage par rapport à la réponse estimée pour une dose nulle d'hormone), le programme calcule l'estimation de la dose (paragraphe II.6.1.) puis l'intervalle de confiance à 95 % (paragraphe II.6.3.). Pour effectuer ce calcul, on suppose que la réponse moyenne est calculée à l'aide de deux répétitions et que le volume de dosage est identique à celui utilisé pour la courbe standard ($\ell=1$, $r_i=2$, $d_1=1$).

Etape 3 : Estimation des doses inconnues

1er cas : $\ell^I = 1$

estimation de la dose inconnue (paragraphe II.6.1.). Calcul d'un intervalle de confiance selon la méthode proposée dans le paragraphe II.6.3.

2ème cas : $\ell^I > 1$

pour chacune des dilutions, le programme revient au cas précédent

```
!dosstat
NOMFICH?sh:cortisol
EXECUTION F4LIB VERSION V11-02 14*44*35*
```

PREPARATION STANDARD

LE NOMBRE TOTAL DE DOSES EST 15
LE NOMBRE TOTAL D OBSERVATIONS EST 64

DOSE	LOG DOSE	COMPTAGES
.00000E+00	DOSE NULLE	2868. 2785. 2849. 2805. 2779. 2588. 2701. 2752.
.20000E-01	-.16990E+01	2615. 2506. 2651. 2498.
.40000E-01	-.13979E+01	2474. 2573. 2378. 2494.
.60000E-01	-.12218E+01	2152. 2307. 2101. 2216.
.80000E-01	-.10969E+01	2016. 2114. 2052. 2030.
.10000E+00	-.10000E+01	1862. 1935. 1800. 1871.
.20000E+00	-.69897E+00	1364. 1412. 1377. 1304.
.40000E+00	-.39794E+00	910. 919. 855. 875.
.60000E+00	-.22185E+00	702. 701. 689. 696.
.80000E+00	-.96910E-01	586. 596. 561. 562.
.10000E+01	.00000E+00	501. 495. 478. 493.
.15000E+01	.17609E+00	392. 358. 399. 394.
.20000E+01	.30103E+00	330. 351. 343. 333.
.40000E+01	.60206E+00	250. 261. 244. 242.
.20000E+05	DOSE INFINIE	131. 135. 134. 133.

PREPARATIONS INCONNUES

IL Y A 118 DOSES INCONNUES A ESTIMER

NUMERO	DILUTION	COMPTAGES
1	1.00	1650. 1750.
2	1.00	1264. 1284.
3	1.00	1843. 1798.
4	1.00	1576. 1475.
5	1.00	1438. 1525.
6	1.00	1412. 1440.
7	1.00	1496. 1500.
8	1.00	298. 293.
9	1.00	316. 342.
10	1.00	293. 331.
11	1.00	338. 359.
12	1.00	433. 424.
13	1.00	478. 518.
.		
.		
.		
.		
.		
.		
.		
.		
105	1.00	1280. 1260.
106	1.00	1938. 1935.
107	1.00	1858. 1920.
108	1.00	1841. 1814.
109	1.00	1865. 1838.
110	1.00	1073. 1081.
111	1.00	1470. 1492.
112	1.00	1932. 1997.
113	1.00	2018. 2017.
114	1.00	2393. 2501.
115	1.00	1483. 1488.
116	1.00	1351. 1388.
117	1.00	1808. 1792.
118	1.00	1988. 1898.
119	1.00	1314. 1328.
120	1.00	2456. 2383.

VOULEZ-VOUS LA NOTICE D UTILISATION DU PROGRAMME?
?oui
Il existe 2 versions : DOSSTAT et DOSSTAT2.
Une notice precise et complete de ces programmes se trouve dans l'annexe C de [Huet S., 1983, l'analyse statistique des dosages radioimmunologiques, these de 3eme cycle, disponible au laboratoire de biometrie du CNRZ].

Ce programme permet d'effectuer l'analyse statistique d'un R.I.A.

On estime d'abord les parametres de la courbe standard puis les doses inconnues.

LE MODELE :

La courbe standard est une sigmoide comprise entre 2 asymptotes D et C .
Si x est le logarithme de la dose z, si U est l'esperance de la reponse pour une dose z,
si V est la variance de la reponse pour une dose z,
Soit F(y)=1/(1+EXP(-y)) avec y=A+Bx

On suppose que U=C+(D-C)*F(y) et V=S2*PHI(H,U)

Une fois que l'utilisateur a choisi son modele, le programme estime les parametres C,D,A,B,S2

Choix possibles :

1- Si la reponse pour une dose infinie d'hormone froide n'a pas ete mesuree, (cela est frequent pour la prolactine par exemple), on peut supposer

-soit que cette reponse est nulle. On a alors U=D*F(y) et le programme estime les parametres D,A,B,et S2.

-soit qu'elle n'est pas nulle et qu'il faut l'estimer. On a alors U=C+(D-C)*F(y) et le programme estime C,D,A,B et S2.

2- Il se peut que la courbe de reponse ne soit pas symetrique. On peut casser la symetrie de la courbe

-soit en ecrivant U=C+(D-C)*(F(y)**GAMMA) avec GAMMA>0

-soit en ecrivant U=C+(D-C)*(1-(1-F(y))**GAMMA) avec GAMMA<0.

Selon la valeur de GAMMA la courbe tend plus ou moins vite vers les asymptotes. Il faut alors introduire la valeur de GAMMA . Si l'utilisateur le desire, le programme fournit une estimation "grossiere" de GAMMA.

3- Il faut choisir la fonction phi qui traduit la variabilite de la variance

-soit la variance est constante V=S2

-soit PHI(H,U)=U**H

-soit PHI(H,U)=U(1+HU)

Pour ces 2 derniers cas il faut introduire la valeur de H. Si l'utilisateur le desire, le programme fournit une estimation de H. Cette estimation de H n'a de sens que si le nombre de reponses observees est grand, et si ces reponses sont bien reparties sur tout l'intervalle des observations.

L'utilisateur peut laisser au programme le choix du modele. Les parametres C,D,A,B et S2 seront alors estimes, la courbe etant supposee symetrique, et la variance des reponses proportionnelle a U.

Une fois les parametres de la courbe standard estimes, le programme fournit la statistique de test d'adequation au modele. Si le test est significatif, il faut essayer d'en decouvrir la cause, puis modifier en consequence les parametres GAMMA ou H. Le tableau des residus reduits et le graphe de la courbe standard sont une aide pour cette recherche.

Si la version DOSSTAT a ete choisie le programme fournit des intervalles de confiance a 95% pour les doses inconnues a l aide du theoreme de Fieller.

Si la version DOSSTAT2 a ete choisie, le programme fournit les limites d'utilisation de la courbe standard (voir chapitre II-6-3) puis les intervalles de confiance a 95% pour les doses inconnues

VOULEZ-VOUS CHOISIR VOUS MEME LES PARAMETRES DE VOTRE MODELE ? (SINON LE PROGRAMME S EN CHARGE)
?non

LE MODELE EST LE SUIVANT

SI X EST LE LOGARITHME DE LA DOSE Z
SI E(U/Z) EST L ESPERANCE DE LA REPONSE POUR UME DOSE Z
SI V(U/Z) EST LA VARIANCE DE LA REPONSE POUR UNE DOSE Z

SOIT F(Y)=1/(1+EXP(-Y)) AVEC Y=A+BX

E(U/Z)=C+(D-C)*F(Y)** 1.0

LA VARIANCE DES REPONSES EST PROPORTIONNELLE A U**H V(U/Z)=S2*U** 1.00

ESTIMATION DES PARAMETRES DE LA COURBE STANDARD

	PARAMETRE ESTIME	ECART TYPE
C	149.48	7.98
D	2831.40	22.86
A	-1.89829	.03392
B	-2.53819	.04633

VARIANCE RESIDUELLE ESTIMEE S2= .191604E+01

TEST D ADEQUATION AU MODELE

SH= 114.963 EST LA SOMME DES CARRES RESIDUELS SA= 67.497 EST LA SOMME DES CARRES INTRA-DOSES
LA STATISTIQUE DE TEST VAUT 3.13 IL FAUT COMPARER CETTE VALEUR A CELLE D UN FISCHER A 11 ET 49 DEGRES DE LIBERTE

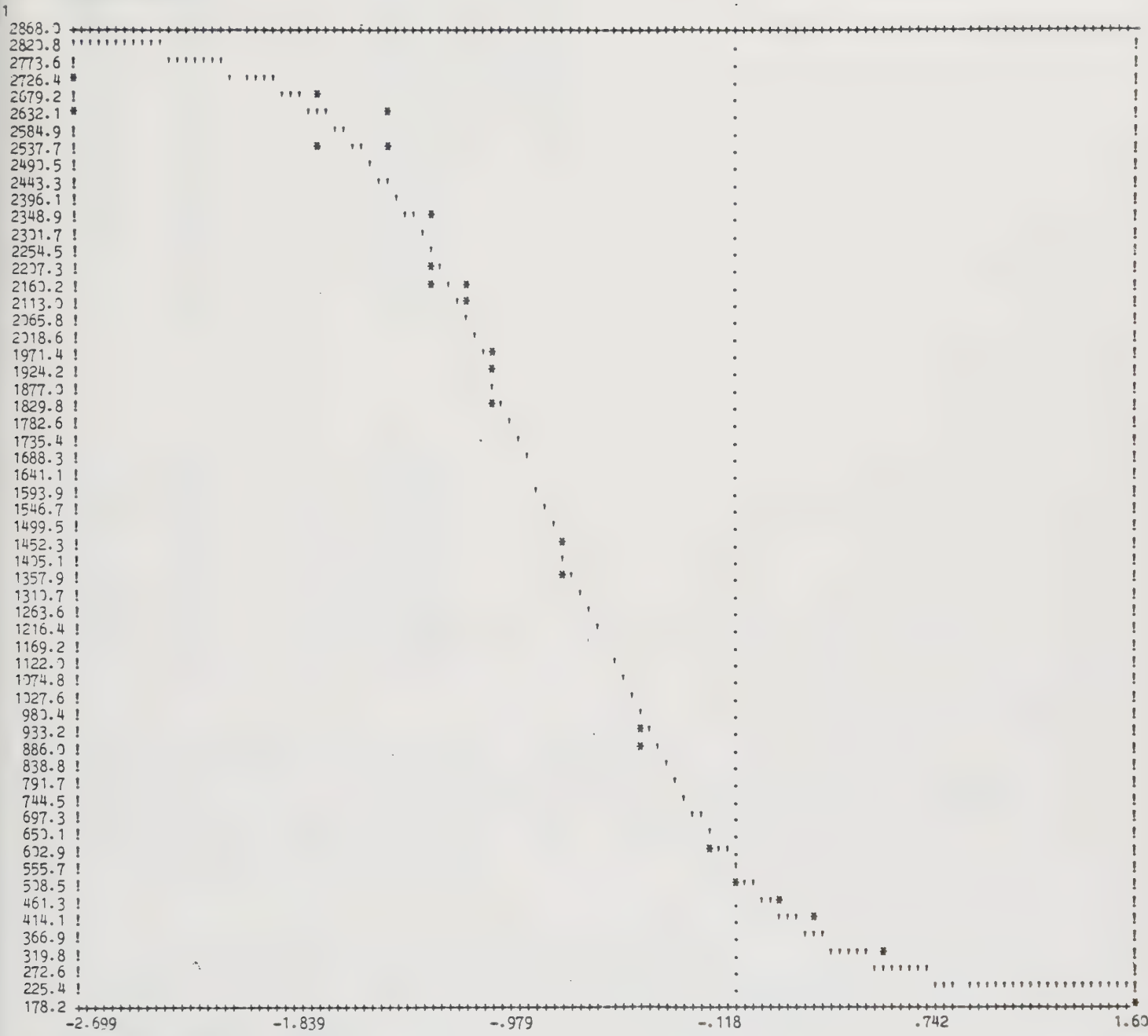
```
*****************************************************************************************************************
*                                                                                                               *
*                                                                                                               *
*                                                                                                               *
*                          AAAAA TTTTT TTTTT EEEEE N   N TTTTT I  OOO  N   N                                    *
*                          A   A   T     T   E     NN  N   T   I O   O NN  N                                    *
*                          AAAAA   T     T   EEE   N N N   T   I O   O N N N                                    *
*                          A   A   T     T   E     N  NN   T   I O   O N  NN                                    *
*                          A   A   T     T   EEEEE N   N   T   I  OOO  N   N                                    *
*                                                                                                               *
*                                                                                                               *
*                                                                                                               *
*          L     EEEEE     M   M  OOO  DDDD  EEEEE L     EEEEE     EEEEE  SSSS TTTTT     FFFFF AAAAA U   U X   X *
*          L     E         MM MM O   O D   D E     L     E         E     S       T       F     A   A U   U  X X  *
*          L     EEE       M M M O   O D   D EEE   L     EEE       EEE    SSS    T       FFF   AAAAA U   U   X   *
*          L     E         M   M O   O D   D E     L     E         E       S   T       F     A   A U   U  X X  *
*          LLLLL EEEEE     M   M  OOO  DDDD  EEEEE LLLLL EEEEE     EEEEE SSSS    T       F     A   A  UUU  X   X *
*                                                                                                               *
*                                                                                                               *
*                                                                                                               *
*          NIVEAU DU TEST :  .0031                                                                              *
*                                                                                                               *
*                                                                                                               *
*****************************************************************************************************************
```

TABLEAU DES RESIDUS REDUITS

DOSE	REPONSE OBSERVEE	REPONSE ESTIMEE	RESIDU	RESIDU REDUIT
NULLE	2868.	2801.40	66.60	.91
NULLE	2785.	2801.40	-16.40	-.22
NULLE	2849.	2801.40	47.60	.65
NULLE	2805.	2801.40	3.60	.05
NULLE	2779.	2801.40	-22.40	-.31
NULLE	2588.	2801.40	-213.40	-2.91
NULLE	2701.	2801.40	-100.40	-1.37
NULLE	2752.	2801.40	-49.40	-.67
.200000E-01	2615.	2583.65	31.35	.45
.200000E-01	2506.	2583.65	-77.65	-1.10
.200000E-01	2651.	2583.65	67.35	.96
.200000E-01	2498.	2583.65	-85.65	-1.22
.400000E-01	2474.	2374.13	99.87	1.48
.400000E-01	2573.	2374.13	198.87	2.95
.400000E-01	2378.	2374.13	3.87	.06
.400000E-01	2494.	2374.13	119.87	1.78
.600000E-01	2152.	2188.95	-36.95	-.57
.600000E-01	2307.	2188.95	118.05	1.82
.600000E-01	2101.	2188.95	-87.95	-1.36
.600000E-01	2216.	2188.95	27.05	.42
.800000E-01	2016.	2027.14	-11.14	-.18
.800000E-01	2114.	2027.14	86.86	1.39
.800000E-01	2052.	2027.14	24.86	.40
.800000E-01	2030.	2027.14	2.86	.05
.100000E+00	1862.	1885.77	-23.77	-.40
.100000E+00	1935.	1885.77	49.23	.82
.100000E+00	1800.	1885.77	-85.77	-1.43
.100000E+00	1871.	1885.77	-14.77	-.25
.200000E+00	1364.	1393.22	-29.22	-.57
.200000E+00	1412.	1393.22	18.78	.36
.200000E+00	1377.	1393.22	-16.22	-.31
.200000E+00	1304.	1393.22	-89.22	-1.73
.400000E+00	910.	922.44	-12.44	-.30
.400000E+00	919.	922.44	-3.44	-.08
.400000E+00	855.	922.44	-67.44	-1.60
.400000E+00	875.	922.44	-47.44	-1.13
.600000E+00	702.	701.88	.12	.00
.600000E+00	701.	701.88	-.88	-.02
.600000E+00	689.	701.88	-12.88	-.35
.600000E+00	696.	701.88	-5.88	-.16
.800000E+00	586.	575.90	10.10	.30
.800000E+00	596.	575.90	20.10	.61
.800000E+00	561.	575.90	-14.90	-.45
.800000E+00	562.	575.90	-13.90	-.42
.100000E+01	501.	495.03	5.97	.19
.100000E+01	495.	495.03	-.03	-.00
.100000E+01	478.	495.03	-17.03	-.55
.100000E+01	493.	495.03	-2.03	-.07
.150000E+01	392.	381.38	10.62	.39
.150000E+01	358.	381.38	-23.38	-.86
.150000E+01	399.	381.38	17.62	.65
.150000E+01	394.	381.38	12.62	.47
.200000E+01	330.	322.47	7.53	.30
.200000E+01	351.	322.47	28.53	1.15
.200000E+01	343.	322.47	20.53	.83
.200000E+01	333.	322.47	10.53	.42
.400000E+01	250.	232.96	17.04	.81
.400000E+01	261.	232.96	28.04	1.33
.400000E+01	244.	232.96	11.04	.52
.400000E+01	242.	232.96	9.04	.43
INFINIE	131.	149.48	-18.48	-1.09
INFINIE	135.	149.48	-14.48	-.86
INFINIE	134.	149.48	-15.48	-.91
INFINIE	133.	149.48	-16.48	-.97

GRAPHE DE LA COURBE STANDARD

AXE HORIZONTAL : LOGARITHME EN BASE 10 DE LA DOSE
AXE VERTICAL : COMPTAGES
+++ : ASYMPTOTE SUPERIEURE D
--- : ASYMPTOTE INFERIEURE C
*** : REPONSES OBSERVEES
''' : COURBE ESTIMEE



NOMBRE DE POINTS DOUBLES = 43

VOULEZ-VOUS REESTIMER LES PARAMETRES DE LA COURBE STANDARD EN CHOISISSANT UN NOUVEA
?oui

VOULEZ-VOUS ESTIMER LA VARIABILITE DE LA VARIANCE?
?oui

MOYENNE	VARIANCE	REPETITION	OBSERVATIONS
2765.9	7921.	8	2868. 2785. 2849. 2805. 2779. 2588. 2701. 2752.
2567.5	5947.	4	2615. 2506. 2651. 2498.
2479.8	6428.	4	2474. 2573. 2378. 2494.
2194.0	7889.	4	2152. 2307. 2101. 2216.
2053.0	1873.	4	2016. 2114. 2052. 2030.
1867.0	3051.	4	1862. 1935. 1800. 1871.
1364.3	2024.	4	1364. 1412. 1377. 1304.
889.8	897.	4	910. 919. 855. 875.
697.0	35.	4	702. 701. 689. 696.
576.3	307.	4	586. 596. 561. 562.
491.8	96.	4	501. 495. 478. 493.
385.8	351.	4	392. 358. 399. 394.
339.3	92.	4	330. 351. 343. 333.
249.3	73.	4	250. 261. 244. 242.
133.3	3.	4	131. 135. 134. 133.
1700.0	5000.	2	1650. 1750.
1274.0	200.	2	1264. 1284.
1820.5	1013.	2	1843. 1798.
1525.5	5101.	2	1576. 1475.
1481.5	3785.	2	1438. 1525.
1426.0	392.	2	1412. 1440.
1498.0	8.	2	1496. 1500.
295.5	13.	2	298. 293.
329.0	338.	2	316. 342.
312.0	722.	2	293. 331.
⋮			
1851.5	365.	2	1865. 1838.
1077.0	32.	2	1073. 1081.
1481.0	242.	2	1470. 1492.
1964.5	2113.	2	1932. 1997.
2017.5	1.	2	2018. 2017.
2447.0	5832.	2	2393. 2501.
1485.5	13.	2	1483. 1488.
1369.5	685.	2	1351. 1388.
1800.0	128.	2	1808. 1792.
1943.0	4050.	2	1988. 1898.
1321.0	98.	2	1314. 1328.
2419.5	2665.	2	2456. 2383.

ESTIMATION DE H : VARIANCE PROPORTIONNNELLE A U**H V(U/Z)=S2*(U**H)

H A ETE ESTIME PAR 1.7 AVEC UN ECART TYPE DE .16 S2 A ETE ESTIME PAR 8526E-02

LA STATISTIQUE DE TEST D ADEQUATION AU MODELE VAUT 31.5 SON ECART TYPE VAUT 19.4

IL N Y A AUCUNE RAISON DE SUPPOSER LE MODELE FAUX

ESTIMATION DE H : VARIANCE PROPORTIONNELLE A U(1+HU) V(U/Z)=S2*U*(1+H*U)

H A ETE ESTIME PAR .18E-02 AVEC UN ECART TYPE DE .47E-03 S2 A ETE ESTIME PAR .2673

LA STATISTIQUE DE TEST D ADEQUATION AU MODELE VAUT 31.3 SON ECART TYPE VAUT 19.4

IL N Y A AUCUNE RAISON DE SUPPOSER LE MODELE FAUX

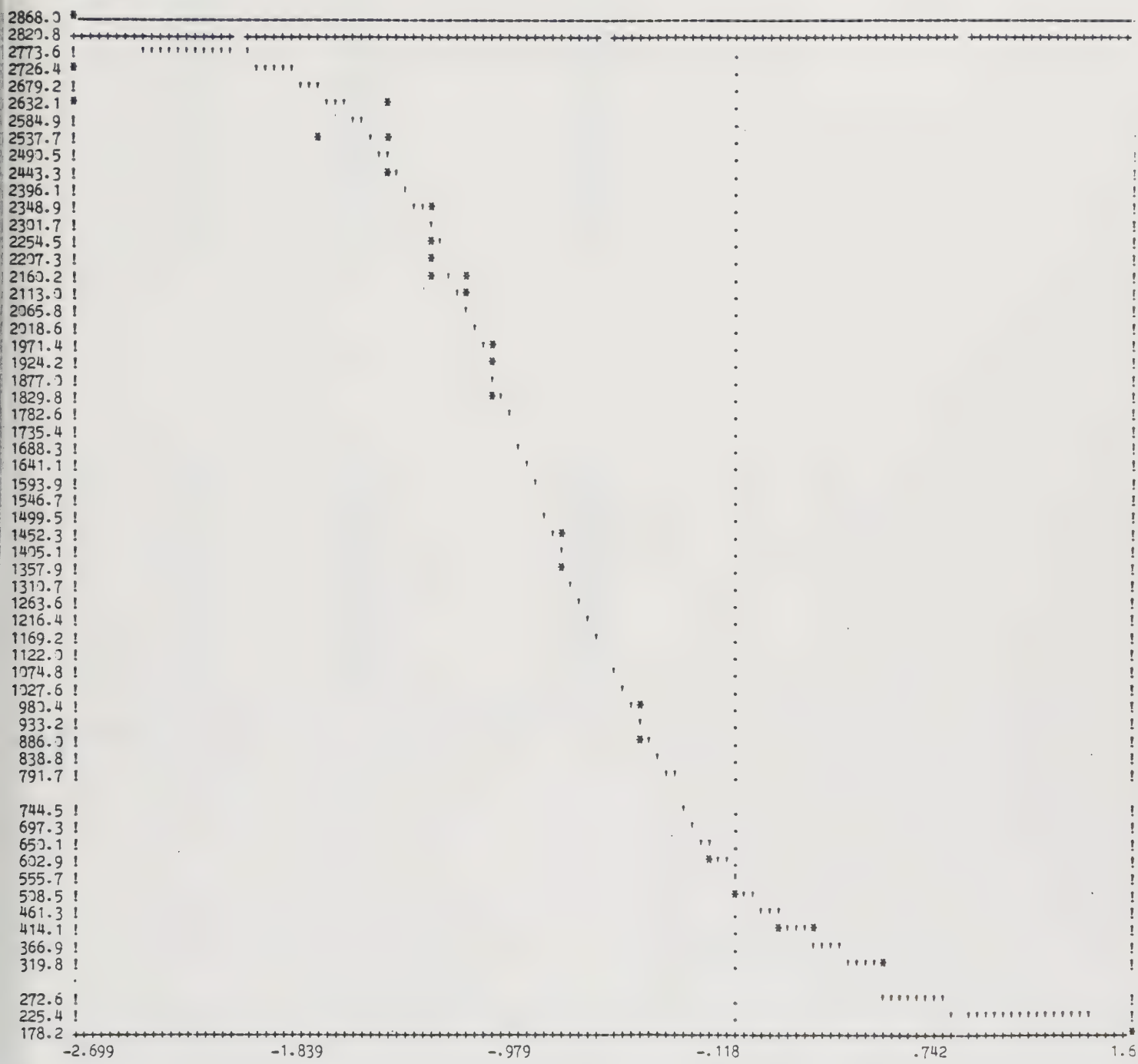
```
  VOULEZ-VOUS UNE ESTIMATION GROSSIERE DE GAMMA?
?oui


  SI GAMMA EST POSITIF,  GAMMA A ETE ESTIME PAR  .62
  L INTERVALLE DE CONFIANCE POUR GAMMA A ETE ESTIME PAR    .5        .7


  OU BIEN SI GAMMA EST NEGATIF,  GAMMA A ETE ESTIME PAR -2.7
  L INTERVALLE DE CONFIANCE POUR GAMMA A ETE ESTIME PAR   -4.       -1.









 LA COURBE DE REPONSE EST-ELLE SUPPOSEE SYMETRIQUE?
?nbn
 QUELLE EST LA VALEUR DE GAMMA?
?.6
 LA VARIANCE DE LA REPONSE EST-ELLE CONSTANTE?
?non
 LA VARIANCE DE LA REPONSE EST-ELLE PROPORTIONELLE A U**H ?
?oui
 QUELLE EST LA VALEUR DE H ?
?1.7



LE MODELE EST LE SUIVANT            SI X EST LE LOGARITHME DE LA DOSE Z
                                    SI E(U/Z) EST L ESPERANCE DE LA REPONSE POUR UME DOSE Z
                                    SI V(U/Z) EST LA VARIANCE DE LA REPONSE POUR UNE DOSE Z

SOIT F(Y)=1/(1+EXP(-Y)) AVEC Y=A+BX

E(U/Z)=C+(D-C)*F(Y)**   .6

LA VARIANCE DES REPONSES EST PROPORTIONNELLE A U**H    V(U/Z)=S2*U** 1.70

 AVEZ-VOUS CORRECTEMENT CHOISI VOTRE MODELE?
?oui



ESTIMATION DES PARAMETRES DE LA COURBE STANDARD

     PARAMETRE ESTIME     ECART TYPE


  C       133.23              2.69
  D      2764.71             21.66
  A        -3.25001              .02432
  B        -3.28375              .03672


VARIANCE RESIDUELLE ESTIMEE  S2=    .778664E-02


TEST D ADEQUATION AU MODELE

 SH=        .467 EST LA SOMME DES CARRES RESIDUELS      SA=        .386 EST LA SOMME DES CARRES INTRA-DOSES
LA STATISTIQUE DE TEST VAUT       .94    IL FAUT COMPARER CETTE VALEUR A CELLE D UN FISCHER A  11 ET  49 DEGRES DE LIBERTE

 IL N Y A AUCUNE RAISON DE SUPPOSER LE MODELE FAUX
```

TABLEAU DES RESIDUS REDUITS

DOSE	REPONSE OBSERVEE	REPONSE ESTIMEE	RESIDU	RESIDU REDUIT
NULLE	2868.	2764.71	103.29	1.39
NULLE	2785.	2764.71	20.29	.27
NULLE	2849.	2764.71	84.29	1.13
NULLE	2805.	2764.71	40.29	.54
NULLE	2779.	2764.71	14.29	.19
NULLE	2588.	2764.71	-176.71	-2.38
NULLE	2701.	2764.71	-63.71	-.86
NULLE	2752.	2764.71	-12.71	-.17
.200000E-01	2615.	2621.99	-6.99	-.10
.200000E-01	2506.	2621.99	-115.99	-1.63
.200000E-01	2651.	2621.99	29.01	.41
.200000E-01	2498.	2621.99	-123.99	-1.75
.400000E-01	2474.	2422.10	51.90	.78
.400000E-01	2573.	2422.10	150.90	2.27
.400000E-01	2378.	2422.10	-44.10	-.66
.400000E-01	2494.	2422.10	71.90	1.08
.600000E-01	2152.	2224.49	-72.49	-1.17
.600000E-01	2307.	2224.49	82.51	1.34
.600000E-01	2101.	2224.49	-123.49	-2.00
.600000E-01	2216.	2224.49	-8.49	-.14
.800000E-01	2016.	2044.95	-28.95	-.50
.800000E-01	2114.	2044.95	69.05	1.20
.800000E-01	2052.	2044.95	7.05	.12
.800000E-01	2030.	2044.95	-14.95	-.26
.100000E+00	1862.	1886.87	-24.87	-.46
.100000E+00	1935.	1886.87	48.13	.90
.100000E+00	1800.	1886.87	-86.87	-1.62
.100000E+00	1871.	1886.87	-15.87	-.30
.200000E+00	1364.	1353.79	10.21	.25
.200000E+00	1412.	1353.79	58.21	1.44
.200000E+00	1377.	1353.79	23.21	.57
.200000E+00	1304.	1353.79	-49.79	-1.23
.400000E+00	910.	889.97	20.03	.71
.400000E+00	919.	889.97	29.03	1.02
.400000E+00	855.	889.97	-34.97	-1.23
.400000E+00	875.	889.97	-14.97	-.53
.600000E+00	702.	686.60	15.40	.68
.600000E+00	701.	686.60	14.40	.63
.600000E+00	689.	686.60	2.40	.11
.600000E+00	696.	686.60	9.40	.41
.800000E+00	586.	572.48	13.52	.69
.800000E+00	596.	572.48	23.52	1.21
.800000E+00	561.	572.48	-11.48	-.59
.800000E+00	562.	572.48	-10.48	-.54
.100000E+01	501.	499.17	1.83	.11
.100000E+01	495.	499.17	-4.17	-.24
.100000E+01	478.	499.17	-21.17	-1.22
.100000E+01	493.	499.17	-6.17	-.36
.150000E+01	392.	394.47	-2.47	-.17
.150000E+01	358.	394.47	-36.47	-2.57
.150000E+01	399.	394.47	4.53	.32
.150000E+01	394.	394.47	-.47	-.03
.200000E+01	330.	338.35	-8.35	-.67
.200000E+01	351.	338.35	12.65	1.02
.200000E+01	343.	338.35	4.65	.37
.200000E+01	333.	338.35	-5.35	-.43
.400000E+01	250.	247.19	2.81	.29
.400000E+01	261.	247.19	13.81	1.45
.400000E+01	244.	247.19	-3.19	-.33
.400000E+01	242.	247.19	-5.19	-.54
INFINIE	131.	133.23	-2.23	-.39
INFINIE	135.	133.23	1.77	.31
INFINIE	134.	133.23	.77	.14
INFINIE	133.	133.23	-.23	-.04

GRAPHE DE LA COURBE STANDARD

AXE HORIZONTAL : LOGARITHME EN BASE 10 DE LA DOSE
AXE VERTICAL : COMPTAGES
+++ : ASYMPTOTE SUPERIEURE D
--- : ASYMPTOTE INFERIEURE C
*** : REPONSES OBSERVEES
''' : COURBE ESTIMEE



JUSQU A MIDI .
NOMBRE DE POINTS DOUBLES = 53

VOULEZ-VOUS REESTIMER LES PARAMETRES DE LA COURBE STANDARD EN CHOISISSANT UN NOUVEAU MODELE AVANT ES
?non

ESTIMATION DES DOSES INCONNUES

PREPARATION	INTERVALLE DE CONFIANCE A 95%		ESTIMATION
1	.1163	.1435	.1279
2	.2058	.2407	.2226
3	.0985	.1211	.1093
4	.1467	.1743	.1600
5	.1555	.1841	.1693
6	.1674	.1975	.1819
7	.1521	.1804	.1657
8	2.3407	2.9776	2.6386
9	1.8957	2.3590	2.1136
10	2.1001	2.6401	2.3535
11	1.7020	2.0976	1.8886
12	1.1828	1.4199	1.2955
13	.9226	1.0929	1.0039
14	1.2464	1.5011	1.3673
15	.5373	.6252	.5795
16	1.0678	1.2743	1.1661
.			
.			
.			
.			
.			
.			
.			
.			
.			
.			
.			
107	.089396	.111095	.099712
108	.0976	.1200	.1083
109	.0943	.1165	.1049
110	.2737	.3179	.2951
111	.1556	.1842	.1694
112	.079941	.100851	.089841
113	.073646	.094080	.083286
114	.029027	.048544	.037552
115	.1547	.1832	.1684
116	.1806	.2123	.1959
117	.1014	.1242	.1123
118	.082572	.103693	.092584
119	.1929	.2261	.2089
120	.031718	.051051	.040257

VOULEZ-VOUS RECOMMENCER L ANALYSE EN CHOISISSANT UN AUTRE MODELE?
?non
STOP 0

```
!dosstat2
NOMFICH?sh:cortisol
EXECUTION F4LIB VERSION V11-02 12*11*36*
```

PREPARATION STANDARD

LE NOMBRE TOTAL DE DOSES EST 15
LE NOMBRE TOTAL D OBSERVATIONS EST 64

DOSE	LOG DOSE	COMPTAGES
.00000E+00	DOSE NULLE	2868. 2785. 2849. 2805. 2779. 2588. 2701. 2752.
.20000E-01	-.16990E+01	2615. 2506. 2651. 2498.
.40000E-01	-.13979E+01	2474. 2573. 2378. 2494.
.60000E-01	-.12218E+01	2152. 2307. 2101. 2216.
.80000E-01	-.10969E+01	2016. 2114. 2052. 2030.
.10000E+00	-.10000E+01	1862. 1935. 1800. 1871.
.20000E+00	-.69897E+00	1364. 1412. 1377. 1304.
.40000E+00	-.39794E+00	910. 919. 855. 875.
.60000E+00	-.22185E+00	702. 701. 689. 696.
.80000E+00	-.96910E-01	586. 596. 561. 562.
.10000E+01	.00000E+00	501. 495. 478. 493.
.15000E+01	.17609E+00	392. 358. 399. 394.
.20000E+01	.30103E+00	330. 351. 343. 333.
.40000E+01	.60206E+00	250. 261. 244. 242.
.20000E+05	DOSE INFINIE	131. 135. 134. 133.

PREPARATIONS INCONNUES

IL Y A 118 DOSES INCONNUES A ESTIMER

NUMERO	DILUTION	COMPTAGES
1	1.00	1650. 1750.
2	1.00	1264. 1284.
3	1.00	1843. 1798.
4	1.00	1576. 1475.
5	1.00	1438. 1525.
6	1.00	1412. 1440.
7	1.00	1496. 1500.
8	1.00	298. 293.
9	1.00	316. 342.
10	1.00	293. 331.
11	1.00	338. 359.
12	1.00	433. 424.
13	1.00	478. 518.
14	1.00	427. 404.
15	1.00	723. 681.
.		
.		
.		
.		
.		
.		
.		
.		
.		
104	1.00	1318. 1293.
105	1.00	1280. 1260.
106	1.00	1938. 1935.
107	1.00	1858. 1920.
108	1.00	1841. 1814.
109	1.00	1865. 1838.
110	1.00	1073. 1081.
111	1.00	1470. 1492.
112	1.00	1932. 1997.
113	1.00	2018. 2017.
114	1.00	2393. 2501.
115	1.00	1483. 1488.
116	1.00	1351. 1388.
117	1.00	1808. 1792.
118	1.00	1988. 1898.
119	1.00	1314. 1328.
120	1.00	2456. 2383.

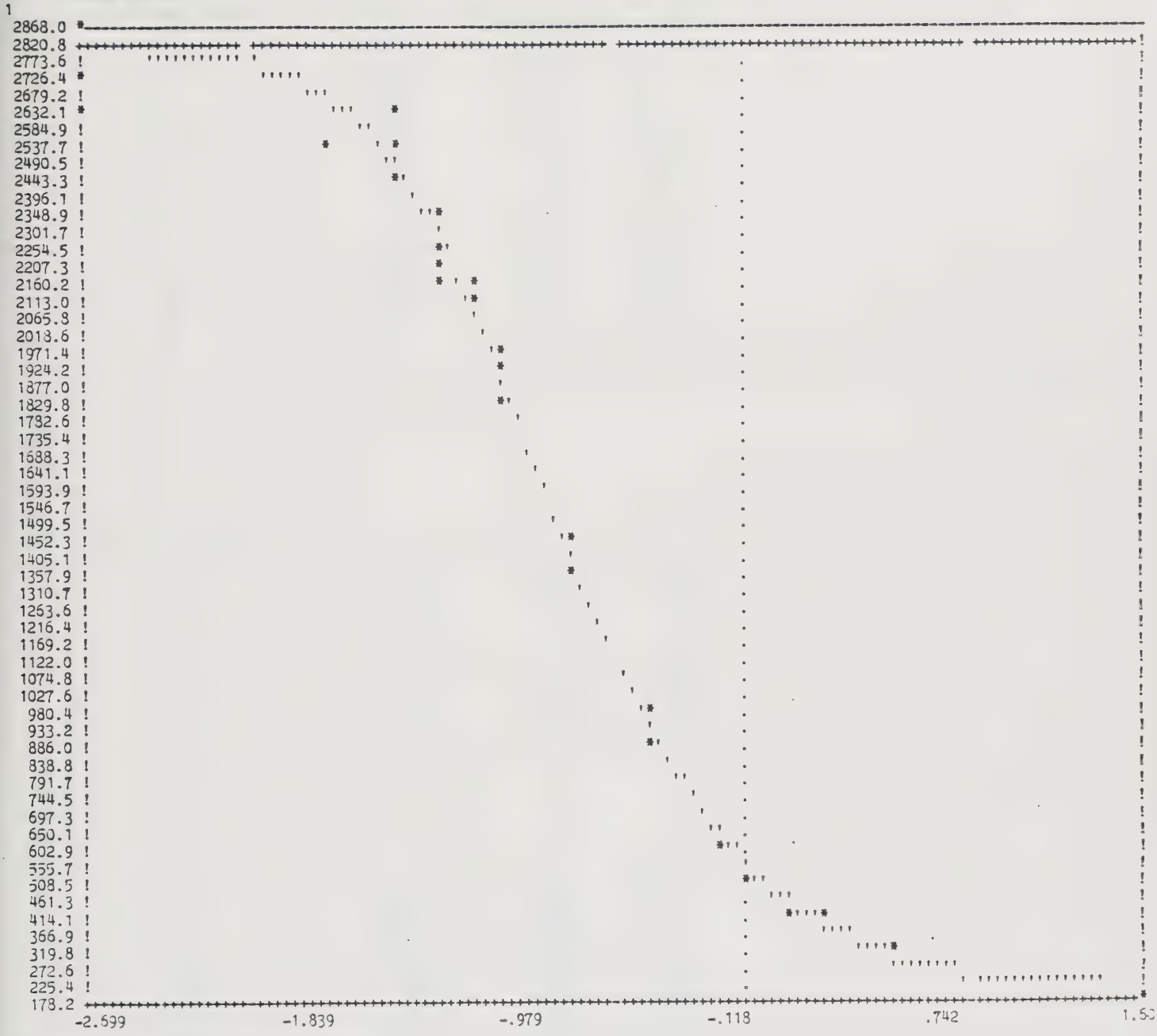
```
 VOULEZ-VOUS LA NOTICE D UTILISATION DU PROGRAMME?
?non

 VOULEZ-VOUS CHOISIR VOUS MEME LES PARAMETRES DE VOTRE MODELE ? (SINON LE PROGRAMME S EN CHARGE)
?oui

  VOULEZ-VOUS ESTIMER LA VARIABILITE DE LA VARIANCE?
?non


  VOULEZ-VOUS UNE ESTIMATION GROSSIERE DE GAMMA?
?non




 LA COURBE DE REPONSE EST-ELLE SUPPOSEE SYMETRIQUE?
?non
 QUELLE EST LA VALEUR DE GAMMA?
?.6
 LA VARIANCE DE LA REPONSE EST-ELLE CONSTANTE?
?non
 LA VARIANCE DE LA REPONSE EST-ELLE PROPORTIONELLE A U**H ?
?oui
 QUELLE EST LA VALEUR DE H ?
?1.7



LE MODELE EST LE SUIVANT            SI X EST LE LOGARITHME DE LA DOSE Z
                                    SI E(U/Z) EST L ESPERANCE DE LA REPONSE POUR UME DOSE Z
                                    SI V(U/Z) EST LA VARIANCE DE LA REPONSE POUR UNE DOSE Z

SOIT F(Y)=1/(1+EXP(-Y)) AVEC Y=A+BX

E(U/Z)=C+(D-C)*F(Y)**   .6

LA VARIANCE DES REPONSES EST PROPORTIONNELLE A U**H    V(U/Z)=S2*U** 1.70

 AVEZ-VOUS CORRECTEMENT CHOISI VOTRE MODELE?
?oui




ESTIMATION DES PARAMETRES DE LA COURBE STANDARD

     PARAMETRE ESTIME     ECART TYPE

  C       133.23              2.69
  D      2764.71             21.66
  A        -3.25001              .02432
  B        -3.28375              .03672


VARIANCE RESIDUELLE ESTIMEE  S2=    .778664E-02


TEST D ADEQUATION AU MODELE

 SH=         .467 EST LA SOMME DES CARRES RESIDUELS      SA=         .386 EST LA SOMME DES CARRES INTRA-DOSES
LA STATISTIQUE DE TEST VAUT        .94    IL FAUT COMPARER CETTE VALEUR A CELLE D UN FISCHER A  11 ET  49 DEGRES DE LIBERTE

 IL N Y A AUCUNE RAISON DE SUPPOSER LE MODELE FAUX
```

TABLEAU DES RESIDUS REDUITS

DOSE	REPONSE OBSERVEE	REPONSE ESTIMEE	RESIDU	RESIDU REDUIT
NULLE	2868.	2764.71	103.29	1.39
NULLE	2785.	2764.71	20.29	.27
NULLE	2849.	2764.71	84.29	1.13
NULLE	2805.	2764.71	40.29	.54
NULLE	2779.	2764.71	14.29	.19
NULLE	2588.	2764.71	-176.71	-2.38
NULLE	2701.	2764.71	-63.71	-.86
NULLE	2752.	2764.71	-12.71	-.17
.200000E-01	2615.	2621.99	-6.99	-.10
.200000E-01	2506.	2621.99	-115.99	-1.63
.200000E-01	2651.	2621.99	29.01	.41
.200000E-01	2498.	2621.99	-123.99	-1.75
.400000E-01	2474.	2422.10	51.90	.78
.400000E-01	2573.	2422.10	150.90	2.27
.400000E-01	2378.	2422.10	-44.10	-.66
.400000E-01	2494.	2422.10	71.90	1.08
.600000E-01	2152.	2224.49	-72.49	-1.17
.600000E-01	2307.	2224.49	82.51	1.34
.600000E-01	2101.	2224.49	-123.49	-2.00
.600000E-01	2216.	2224.49	-8.49	-.14
.800000E-01	2016.	2044.95	-28.95	-.50
.800000E-01	2114.	2044.95	69.05	1.20
.800000E-01	2052.	2044.95	7.05	.12
.800000E-01	2030.	2044.95	-14.95	-.26
.100000E+00	1862.	1886.87	-24.87	-.46
.100000E+00	1935.	1886.87	48.13	.90
.100000E+00	1800.	1886.87	-86.87	-1.62
.100000E+00	1871.	1886.87	-15.87	-.30
.200000E+00	1364.	1353.79	10.21	.25
.200000E+00	1412.	1353.79	58.21	1.44
.200000E+00	1377.	1353.79	23.21	.57
.200000E+00	1304.	1353.79	-49.79	-1.23
.400000E+00	910.	889.97	20.03	.71
.400000E+00	919.	889.97	29.03	1.02
.400000E+00	855.	889.97	-34.97	-1.23
.400000E+00	875.	889.97	-14.97	-.53
.600000E+00	702.	686.60	15.40	.68
.600000E+00	701.	686.60	14.40	.63
.600000E+00	689.	686.60	2.40	.11
.600000E+00	696.	686.60	9.40	.41
.800000E+00	586.	572.48	13.52	.69
.800000E+00	596.	572.48	23.52	1.21
.800000E+00	561.	572.48	-11.48	-.59
.800000E+00	562.	572.48	-10.48	-.54
.100000E+01	501.	499.17	1.83	.11
.100000E+01	495.	499.17	-4.17	-.24
.100000E+01	478.	499.17	-21.17	-1.22
.100000E+01	493.	499.17	-6.17	-.36
.150000E+01	392.	394.47	-2.47	-.17
.150000E+01	358.	394.47	-36.47	-2.57
.150000E+01	399.	394.47	4.53	.32
.150000E+01	394.	394.47	-.47	-.03
.200000E+01	330.	338.35	-8.35	-.67
.200000E+01	351.	338.35	12.65	1.02
.200000E+01	343.	338.35	4.65	.37
.200000E+01	333.	338.35	-5.35	-.43
.400000E+01	250.	247.19	2.81	.29
.400000E+01	261.	247.19	13.81	1.45
.400000E+01	244.	247.19	-3.19	-.33
.400000E+01	242.	247.19	-5.19	-.54
INFINIE	131.	133.23	-2.23	-.39
INFINIE	135.	133.23	1.77	.31
INFINIE	134.	133.23	.77	.14
INFINIE	133.	133.23	-.23	-.04

GRAPHE DE LA COURBE STANDARD

AXE HORIZONTAL : LOGARITHME EN BASE 10 DE LA DOSE
AXE VERTICAL : COMPTAGES
+++ : ASYMPTOTE SUPERIEURE D
--- : ASYMPTOTE INFERIEURE C
*** : REPONSES OBSERVEES
''' : COURBE ESTIMEE



NOMBRE DE POINTS DOUBLES = 53

DETERMINATION DES LIMITES D UTILISATION DE LA COURBE STANDARD

LA PLUS FAIBLE DOSE DE LA COURBE STANDARD VAUT .200E-01 (SOIT 94.8% DE D)
LA PLUS FORTE DOSE VAUT 4.0

LA COURBE EST UTILISABLE POUR LES DOSES SUPERIEURES A .257E-01 (SOIT 92.8% DE D) ET INFERIEURES A 4.0

INTERVALLE DE CONFIANCE CALCULE EN FONCTION DU POURCENTAGE DE REPONSE OBSERVE
(LE NOMBRE DE REPETITIONS EST SUPPOSE EGAL A 2)

POURCENTAGE DE REPONSE OBSERVE	DOSE ESTIMEE	INTERVALLE DE CONFIANCE CALCULE A 95%	
5	154.79	LA REPONSE SORT DE L INTERVALLE D ETUDE DE LA COURBE STANDARD	
10	3.0566	2.5091	3.8635
15	1.3720	1.1856	1.6167
20	.8460	.7438	.9745
25	.5938	.5258	.6771
30	.4467	.3965	.5071
35	.3505	.3109	.3974
40	.2826	.2499	.3209
45	.2319	.2042	.2642
50	.1926	.1684	.2205
55	.1610	.1395	.1857
60	.1348	.1154	.1570
65	.1127	.0949	.1329
70	.093580	.076990	.112250
75	.076652	.060887	.094155
80	.061347	.045970	.078026
85	.047146	.031511	.063377
90	.033498	.016179	.049793
95	.019524	LA REPONSE SORT DE L INTERVALLE D ETUDE DE LA COURBE STANDARD	

VOULEZ-VOUS REESTIMER LES PARAMETRES DE LA COURBE STANDARD EN CHOISISSANT UN NOUVEAU MODELE AVANT ESTIMATION DES DOSES INCONNUES ?
?non

ESTIMATION DES DOSES INCONNUES

NUMERO DE PREPARATION	DOSE ESTIMEE	INTERVALLE DE CONFIANCE CALCULE A 95%	
1	.1279	.1025	.1585
2	.2226	.1957	.2539
3	.1093	.0944	.1258
4	.1600	.1437	.1778
5	.1693	.1537	.1862
6	.1819	.1666	.1984
7	.1657	.1521	.1803
8	2.6386	2.3659	2.9704
9	2.1137	1.9220	2.3405
10	2.3535	2.1373	2.6099
11	1.8886	1.7347	2.0673
.			
.			
.			
.			
.			
.			
.			
104	.2133	.2036	.2231
105	.2238	.2138	.2339
106	.093424	.087594	.099333
107	.099712	.093658	.105834
108	.1083	.1019	.1147
109	.1049	.0986	.1112
110	.2951	.2831	.3071
111	.1694	.1610	.1778
112	.089841	.084162	.095601
113	.083287	.077845	.088821
114	.037552	.032694	.042171
115	.1684	.1601	.1768
116	.1959	.1867	.2050
117	.1123	.1058	.1188
118	.092585	.086828	.098416
119	.2089	.1994	.2185
120	.040257	.035530	.044830

VOULEZ-VOUS RECOMMENCER L ANALYSE EN CHOISISSANT UN AUTRE MODELE?
?non
STOP 0

ANNEXE D

Calcul des extrêma de $F^{-1}(\psi)$ sur le contour de $\mathcal{R}_{1\alpha}$, région de confiance approchée pour ψ.

Ce calcul est très simple à programmer ; il suffit comme le propose ROSS (1978) d'utiliser la méthode des multiplicateurs de Lagrange de la façon suivante. Notons W^{-1} la matrice des dérivées secondes de la vraisemblance en Ψ

$$W^{-1} = \frac{\partial^2 \ell n\, L}{\partial\, \psi^2}(\bar{\Psi}) \quad \text{et posons } c = \chi^2_\alpha\,(5).$$

Nous allons minimiser ou maximiser

$F^{-1}(\psi) - \lambda\left[(\psi-\Psi)'\ W^{-1}(\psi-\Psi) - c\right]$ par rapport à λ et ψ. Ceci nous conduit à résoudre le système d'équations

$$\begin{cases} \dfrac{\partial F^{-1}}{\partial \psi}(\psi) = \lambda\ W^{-1}\ (\psi - \Psi) \\ (\psi-\Psi)'\ W^{-1}\ (\psi-\Psi) = c \end{cases}$$

Ce système est résolu itérativement. Soit ψ_o une valeur initiale pour ψ. Par exemple ψ_o est le vecteur des valeurs calculées pour U^I et Θ. D'après le système d'équations ci-dessus nous calculons $\dfrac{\partial F^{-1}}{\partial \psi}(\psi_o)$ et nous en déduisons

$$\lambda_o^2 = \left(\frac{\partial F^{-1}}{\partial \psi}(\psi_o)\right)'\ W\ \frac{\partial F^{-1}}{\partial \psi}(\psi_o)/c$$

et $\Psi_1 = \Psi + \dfrac{1}{\lambda_o} W\ \dfrac{\partial F^{-1}}{\partial \psi}(\psi_o)$

(λ_o étant la racine positive ou négative de λ_o^2 selon que nous cherchons le

maximum ou la minimum de F^{-1}).

Or F^{-1} et les dérivées de F^{-1} ne sont définies que pour les valeurs de ψ vérifiant $c < \mu^I < d$. (En effet, la courbe de réponse prend ses valeurs entre les deux asymptotes c et d et $F^{-1}(\mu^I,\theta) = (\ell n\ (d-\mu^I)/(\mu^I-c))-\alpha)\ /\ \beta)$

Trois cas peuvent alors se présenter :

1er cas : F^{-1} n'est pas définie en Ψ_1. L'itération est alors stoppée et l'intervalle de confiance pour x^I a une limite supérieure ou inférieure infinie.

2ème cas : F^{-1} est définie en Ψ_1 et l'écart entre Ψ_1 et ψ_o est très faible. Alors $F^{-1}(\Psi_1)$ est l'extrêmum cherché de F^{-1} sur le contour de $\mathcal{R}_{1\alpha}$. C'est-à-dire, si nous cherchions le maximum de F^{-1}, $X_{sup} = F^{-1}(\Psi_1)$, X_{sup} étant la limite supérieure de x_I.

3ème cas : F^{-1} est définie en Ψ_1 mais l'écart entre Ψ_1 et ψ_o est grand. Nous calculons

$$\lambda_1^2 = \left(\frac{\partial F^{-1}}{\partial \psi}(\Psi_1)\right)' W \frac{\partial F^{-1}}{\partial \psi}(\Psi_1) \ /\ c$$

Nous en déduisons Ψ_2 et commençons une nouvelle itération.

C O N C L U S I O N

-=-=-

Ainsi, le travail qui a été fait pour répondre aux questions qui nous étaient posées, nous conduit à formuler de nouvelles interrogations.

1) Pourquoi la courbe logistique ? Sans doute parce que c'est la plus simple et la plus flexible (lorsqu'on modifie les valeurs des paramètres) des courbes sigmoïdes ; elle présente un point d'inflexion et un comportement asymptotique du type cherché. Mais il s'agit d'une démarche évidemment mécaniste qui contient en elle-même de nombreuses limitations. Cependant, la possibilité d'augmenter le nombre de paramètres dans ces courbes pourrait sembler une manière de dépasser cette limitation initiale.

* en effet, en augmentant le nombre de paramètres, nous améliorons la qualité de l'ajustement, nous réduisons la variance résiduelle. En fait la réponse est plus complexe.

* U étant la réponse pour une dose x, écrivons

$$U = F(x,\theta) + E$$

F appartient à une certaine famille de fonctions, θ est un vecteur de dimension p où p est le nombre de paramètres inconnus à estimer.

Supposons que Θ_p soit l'estimateur du maximum de vraisemblance de θ ; l'erreur E est d'espérance nulle et de variance constante σ^2. Plus le nombre de paramètres p sera grand, plus la valeur de la vraisemblance en Θ_p sera grande, et plus la somme des carrés résiduelle sera faible. A la limite, s'il y a autant de paramètres que d'observations la somme des carrés résiduelle est nulle, et l'estimation de σ^2 n'est pas possible. Mais en contre-partie, l'estimation de la variance de $F(x,\Theta_p)$ augmente.

Pour illustrer cela, considérons par exemple le cas très simple de la régression polynomiale.

Supposons que $U_i = F(x_i,\theta) + E_i$, pour i variant de 1 à n où $F(x_i,\theta)$ est un polynôme en x_i de degré p-1 :

$$F(x_i,\theta) = \sum_{\ell=0}^{p-1} \theta_\ell \, \phi_\ell \, (x_i)$$

où $\phi_o(x),\ldots,\phi_{p-1}(x)$ sont p polynômes orthogonaux de degré respectif $0,\ldots,p-1$.

Soit $\hat{\theta}_\ell$ l'estimateur du maximum de vraisemblance de θ

$\hat{\theta}_\ell = \sum_{i=1}^{n} \phi_\ell(x_i)\, U_i$, pour ℓ variant de 0 à p-1.

Les $\hat{\theta}_\ell$ sont de même variance, $\mathrm{Var}\,\hat{\theta}_\ell = \sigma^2$ et sont non corrélés.

La somme des carrés résiduelle, SC^2_{p-1} vaut

$$SC^2_{p-1} = \sum_{i=1}^{n} \left(U_i - \sum_{\ell=0}^{p-1} \hat{\theta}_\ell\, \phi_\ell(x_i)\right)^2.$$ $F(x,\theta)$ est estimé par

$\sum_{\ell=0}^{p-1} \hat{\theta}_\ell\, \phi_\ell(x)$ dont la variance vaut $\sigma^2 \sum_{\ell=0}^{p-1} \phi_\ell(x)^2$. Si nous supposons

maintenant que $F(x,\theta)$ est un polynôme en x de degré p, alors $SC^2_p = SC^2_{p-1} - \hat{\theta}^2_p$ et la variance de l'estimateur de $F(x,\theta)$ vaut $\sigma^2 \sum_{\ell=0}^{p-1} \phi_\ell(x)^2 + \sigma^2\, \phi_p(x)^2$.

Ainsi, plus le nombre de paramètres p augmente plus la vraisemblance en Θ_p augmente, mais plus la qualité prédictive du modèle diminue.

Donc quand nous estimerons la dose inconnue à l'aide de la courbe standard estimée, nous pouvons penser que plus le nombre de paramètres sera grand, plus l'estimation de la dose inconnue sera imprécise (plus la prédiction sera mauvaise). Quand nous voudrons comparer deux modèles, il ne suffira donc pas de comparer les vraisemblances et de choisir celui pour lequel la vraisemblance est la plus grande. Il serait intéressant de proposer des critères qui prennent en compte d'une part un bon ajustement des observations (ceci est d'autant mieux réalisé que la vraisemblance est élevée) et d'autre part une bonne qualité de la prédiction de la dose inconnue. Dans certains domaines de la statistique, notamment l'étude du choix d'un modèle pour la modélisation d'une série chronologique (en vue ou non d'une prédiction), dans le choix de modèles génétiques, de telles recherches ont déjà été menées (AKAIKE 1973). Nous pensons que dans le domaine des dosages biologiques, il pourrait s'agir d'une recherche intéressante.

Ces méthodes reviennent à maximiser non pas la vraisemblance $\mathcal{V}(U,\theta,\theta\in\mathcal{O}_p)$

mais une forme pénalisée de la vraisemblance,

$\mathcal{V}(U,\theta,\ \theta\in \mathcal{O}_p) - \Phi(n,p)$, où n est la dimension de U (le nombre d'observations) et p celle de θ. Le maximum est calculé en p et en $\theta \in \mathcal{O}_p$. La fonction de pénalisation Φ est une fonction croissante de p et décroissante de n. Un exemple courant est celui où $\Phi(n,p) = 2p$ qui conduit à des résultats assez intéressants pour une valeur de n de l'ordre de 30 à 60, mais qui conduit également à des estimateurs non consistants de p.

La vraisemblance pénalisée passe par un maximum pour $(\hat{p},\hat{\theta} \in \mathcal{O}_{\hat{p}})$. $\hat{p}$ est l'ordre identifié du modèle, par un principe dit de parcimonie qui en limitant le nombre de paramètres cherche un compromis entre ajustement et prédiction.

2) robustesse du modèle, à l'estimation de la variabilité de la variance.

Rappelons le modèle :

$$U = F(x,\theta) + E$$

Supposons que l'erreur E est d'espérance nulle, et de variance proportionnelle à $\phi_h(\mu)$ où μ est l'espérance de U en une dose x.

Au paragraphe II.4., nous avons exposé une méthode d'estimation de h pour une fonction ϕ donnée. Les propriétés d'un tel estimateur sont totalement inconnues. JEWELL et RAAB (1981) restreignent leur étude à la consistance de σ^2 dans le cas où $Var(U) = \sigma^2 \mu^2$.

Toutefois, FINNEY et PHILIPPS (1977), FINNEY (1976) assurent que les valeurs des estimateurs des paramètres de la courbe standard et donc les valeurs des estimateurs de la dose inconnue sont très peu modifiées par des variations de h.

Une étude plus précise des raisons de cette robustesse serait sans doute très intéressante.

3) maximum de vraisemblance ou moindres carrés.

Nous avons vu dans l'annexe B, que nous avons deux façons différentes d'estimer les paramètres de la courbe standard.

- chercher c, d, α, β et σ^2 qui minimisent l'opposé du logarithme de la vraisemblance.

$$2\mathcal{L} = \sum_{i=1}^{k} n_i \text{ Log } 2\pi\sigma^2 + \sum_i n_i \ln \phi_h(\mu_i) + \sum_{i,j} \frac{(U_{ij} - \mu_i)^2}{\sigma^2 \phi_h(\mu_i)}$$

avec $$\mu_i = c + \frac{d-c}{1+\exp-(\alpha+\beta x_i)}$$

- chercher c, d, α, et β qui minimisent la somme des carrés $\mathcal{S}$

$$\mathcal{S} = \sum_{i,j} \frac{(U_{ij} - \mu_i)^2}{\phi_h(\mu_i)}$$

Si $\phi_h(\mu_i)$ est constant et vaut 1 pour toutes les valeurs de μ_i, les deux méthodes sont équivalentes. Sinon, l'affirmation n'est pas vraie sauf peut être d'un point de vue asymptotique (c'est-à-dire pour de grandes valeurs de $n=\Sigma n_i$), et dans ce cas, il serait intéressant de la démontrer. De plus, quelles sont les propriétés de ces estimateurs ? Que se passe-t-il "à distance finie", c'est-à-dire pour une valeur de n fixée ? Que vaut le biais des estimateurs ? Quelle est leur loi ? Leur variance est-elle correctement estimée ?

Des problèmes mathématiques assez complexes nécessitant des développements théoriques et des simulations restent donc posés pour apprécier pleinement la validité des techniques usuelles.

B I B L I O G R A P H I E

-=-=-=-=-

** B I B L I O G R A P H I E **

AKAIKE, H. (1973).

Maximum likelihood identification of Gaussian autoregressive moving average models, BIOMETRIKA, 60, 255-265.

BATES, D.M., WATTS, D.G. (1980).

Relative curvature measures of nonlinearity. JOURNAL OF THE ROYAL STATISTICAL SOCIETY. Series B, 42(1), 1-25.

BEALE, E.M.L. (1960).

Confidence regions in non-linear estimation. JOURNAL OF THE ROYAL STATISTICAL SOCIETY. Series B, 22, 41-76.

BARD, Y. (1974).

Nonlinear parameter estimation. Academic Press. New-York.

BOX, M.J. (1968).

The occurence of replications in optimal designs of experiments to estimate parameters in non-linear models. J. ROY. STATIST. SOC. Series B, 30, 290-302.

CHANG, P.C., RUBIN, R.T., MIMI YU. (1975).

Optimal statistical design of radioimmunoassays and competitive protein-binding assays. ENDOCRINOLOGY, 94(4), 973-981.

CORNFIELD, J. (1964).

Comparative bio assays and the role of parallelism. J. PHARMAC. EXP. THER., 144, 143-149.

DACUNHA-CASTELLE, D., DUFLO, M. (1982).

Probabilites et statistiques. Masson, Paris.

DAGNELIE, P. (1969).

Theorie et methodes statistiques. Volumes I et II. J. Duculot, S.A., GEMBLOUX.

ELASHOFF, J.D. (1981).

Repeated measures bioassay with correlated errors and heterogeneous variances : a Monte Carlo study. BIOMETRICS, 37, 475-482.

FEDOROV, V.V. (1972).

Theory of optimal experiments (Birnbaum Z.W., Lukacs E., ed) probability and mathematical statistics. a series of monographs and textbooks. Academic Press, Ney-York, London.

FINNEY, D.J. (1976).

Radioligand assay. BIOMETRICS, 32, 721-740.

FINNEY, D.J. (1977).

Bioassay. In Mathematics and the life sciences. (Levin S., ed). Lecture Notes in Biomathematics, 18, Springer-Verlag.

FINNEY, D.J., PHILIPPS, P. (1977).

The form and estimation of a variance function, with particular reference to radioimmunoassay. THE JOURNAL OF THE ROYAL STATISTICAL SOCIETY. Series C, 26(3), 312-320.

FINNEY, D.J. (1978).

Stastical method in biological assay. Charles Griffin & Co Ltd. London & High Wycombe.

FINNEY, D.J. (1982).

Communication personnelle.

HARWELL (1981).

Harwell subroutine library. A catalogue of subroutines. United Kingdom Atomic Energy Authority. Oxfordshire.

HEALY, M.J.R. (1972).

Statistical analysis of radioimmunoassay data. BIOCHEM. JOURNAL, 130, 207-210.

HUET, S. (1978).

Dosage radioimmunologique. Analyse des resultats. Rapport de DEA.

JEWELL, N.P., RAAB, G.M. (1981).

Difficulties in obtaining consistent estimators of variance parameters. BIOMETRIKA, 68(1), 221-226.

JUPP, D.L.B. (1978).

Approximation to data by splines with free knots. SIAM. J. NUMER. ANAL., 15(2), 328-343.

KENDALL, M.G., STUART, A. (1962).

The advanced theory of statistics. Volumes I, II et III. Charles Griffin et Co. Ltd. London,

LEFORT, G. (1967).

Mathematiques pour les sciences biologiques et agronomiques. Librairie Armand Colin, Paris.

MARQUARDT, D.W. (1963).

An algorithm for least squares estimation of non linear parameters J. SOC. INDUST. APPL. MATH., 11, 431-441.

METZLER, C.M., GARY, E.L., AMBER, M.E.J. (1976).

A users manual for NONLIN and associated programs. Report : research biostatistics. The Upjohn Company Kalamazoo, Michigan 49001.

PEARSON, E.S., HARTLEY, H.O. (1970).

Biometrika tables for statisticians. Volume I. At the University Press. Cambridge.

RAAB, G.M., THOMPSON, R., Mc KENZIE, I. (1980).

Variance function estimation for Immunoassays. Computer programs in Biomedicine, 12, 111-120.

RAAB, G.M. (1981).

Estimation of a variance function with application to Immunoassay APPLIED STATISTICS, 30(1), 32-40.

RAWLINS, T.G.R., YRJONEN, T. (1978).

Calculation of RIA results using the spline function. International Laboratory. Nov./Dec.

RODBARD, D., LEWALD, J.E. (1970).

Computer analysis of radioligand assay and radioimmunoassay data. ACTA ENDOCRINOL., 64, SUPPL. 147, 79-103.

RODBARD, D., FRAZIER, G,R, (1975).

Statistical analysis of radioligand assay data. METHOD IN ENZYMOLOGY, 37, 3-22.

RODBARD, D., LENOX, R.H., WRAY, H.L., RAMSETH, D. (1976).

Statistical characterization of the random errors in the radioim munoassay dose-response variable CLINICAL CHEMISTRY,

ROSS, G.J.S. (1978).

Exact and approximate confidence regions for functions of parameters in Non-linear models, in COMPSTAT 1978 (Corsten L.C.A & HERMANS J. Editors), Physica-Verlag, Wien, 110-116).

SANDEL, P., VOGT, W. (1977).

Performance of various mathematical methods for calculation of radioimmunoassay results. International Symposium on radioimmunoassay and related procedures in medicine, IAEA/WHO, Berlin.

SNEDECOR, G.W., COCHRAN, W.G. (1971).

Methodes statistiques. ACTA, Paris.

TOMASSONE, R., LESQUOY, E., MILLIER, C. (1982).

La regression : regards nouveaux sur une methode ancienne. A.S.A., Masson-INRA, Paris.

VILA, J.P. (1976).

M.L.P. Presentation, utilisation. Note interne du Laboratoire de Biometrie du C.N.R.Z.

VOLUND, A. (1978).

Application of the four parameter logistic model to bioassay : comparison with slope ratio and parallel line models. BIOMETRICS, 34, 357-365.

WOLD, S. (1974).

Spline functions in data analysis. TECHNOMETRICS, 16(1), 1-11.

S U M M A R Y

Statistical analysis of radioimmunological assays is usually founded on statistical assumption which are unknown or forgotten, often false, seldom checked.

A critical examination of usual practices is made, after which, some new methods for statistical analysis are set up. These methods have often been defined elsewhere, but new one have been worked out and comply well with biologists queries.

A conversationnel program, very easy to use, is set up.

Key-words : biological assay - radioimmunological assay - linear model - non-linear model.